

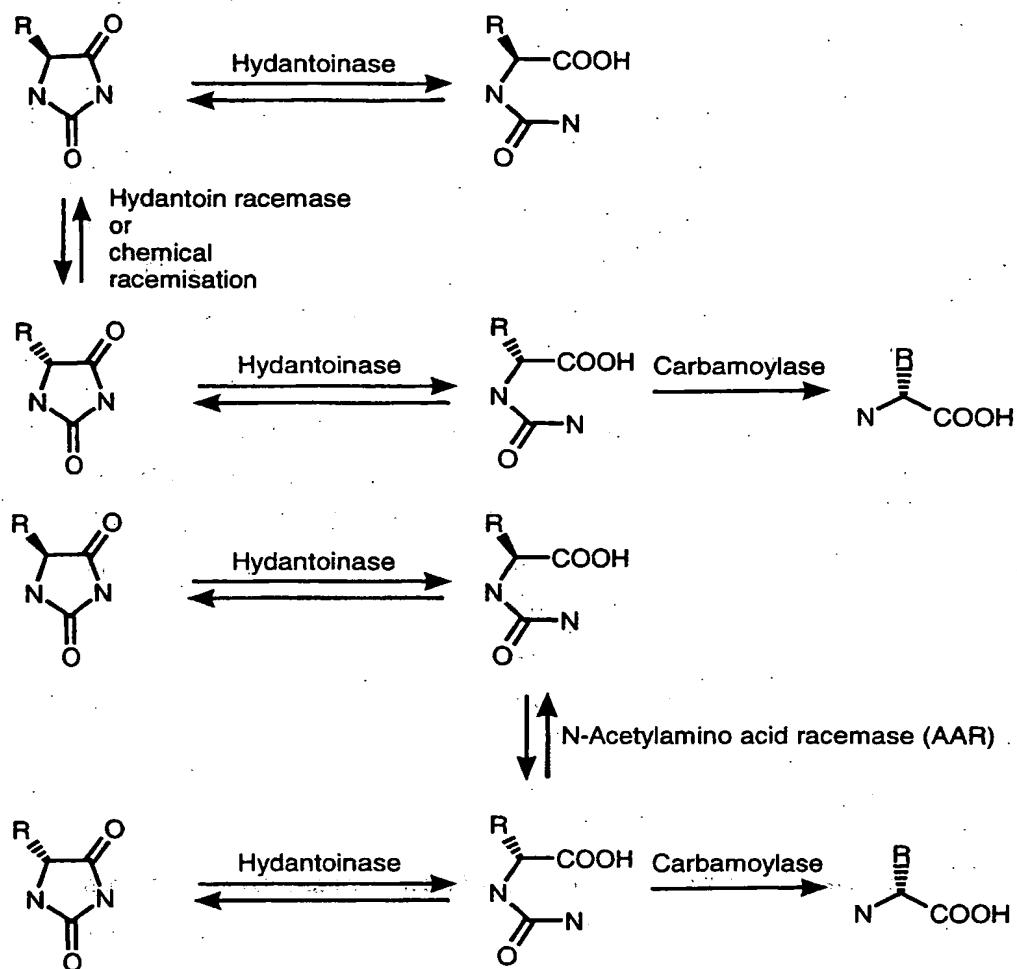
Screening process for hydantoin racemases

The present invention relates to a screening process for the detection of improved hydantoin racemases, to novel hydantoin racemases themselves and to their use in the 5 preparation of N-carbamoyl-amino acids.

These optically active compounds are compounds that are frequently used in organic synthesis for the preparation of, for example, active ingredients having biological activity. They are also present in chiral auxiliaries, for 10 example in the form of the amino alcohols (Evans reagents).

The enzymatic hydrolysis of 5-substituted hydantoins to N-carbamoyl-amino acids and the further reaction thereof to the corresponding enantiomerically enriched amino acids is a standard method in organic chemistry ("Enzyme Catalysis 15 in Organic Synthesis", Eds.: Drauz, Waldmann, VCH, 1st and 2nd Ed.). The enantiodifferentiation can be carried out either at the stage of the hydantoin hydrolysis by means of hydantoinases or alternatively during the cleavage of the N-carbamoyl-amino acids by means of enantioselective 20 carbamoylases. Because the enzymes each convert only one optical antipode of the corresponding compound, it is attempted to racemise the other in the mixture (*in situ*) in order to ensure the complete conversion of the hydantoin, which can readily be prepared racemically, into the 25 corresponding enantiomerically enriched amino acid. The racemisation can proceed either at the stage of the hydantoins by means of chemical (base, acid, elevated temperature) or enzymatic processes or alternatively at the stage of the N-carbamoyl-amino acids by means of, for 30 example, acetyl amino acid racemases (DE10050124). By its nature, the latter variant is only successful if enantioselective carbamoylases are used. The following scheme illustrates this fact.

Scheme 1:



For aromatic substrates, the rate of the chemical racemisation of the hydantoins, as shown in Table 1, is sufficiently high to ensure high space-time yields for the preparation of amino acids by the hydantoinase process. For 10 aliphatic hydantoins, such as isobutyl-, methyl- and isopropyl-hydantoin, however, the racemisation represents a considerable bottleneck in the synthesis of aliphatic amino acids.

Table 1: Racemisation constants of hydantoins at 40°C, pH 8.5 determined by initial rates according to a first-order reaction ($-k_{rac} = \ln([a]/[a_0])$) from: Hydrolysis and Formation of Hydantoins (Chpt. B 2.4). Syldatk, C. and Pietzsch, M. In: Enzyme catalysis in organic synthesis (Eds.: K. Drauz & H. Waldmann), VCH, 1st and 2nd Ed.).

5'-substituent	k_{rac} (h ⁻¹)	$t_{1/2}$ (h)
Phenyl	2.59	0.27
Methylthioethyl	0.12	5.82
Isobutyl	0.032	21.42
Methyl	0.02	33.98
Isopropyl	0.012	55.90

This problem manifests itself, for example, in the preparation, described in EP759475, of enantiomerically enriched tert.-butylhydantoin by means of the hydantoinase process. In this case, the complete conversion of 32 mM tert.-butylhydantoin with 1.5 kU R-hydantoinase required 8 days at pH 8.5 and 4 days at pH 9.5. The low space-time yield is in fact caused by the only slow chemical racemisation of tert.-butylhydantoin ($k_{rac} = 0.009$ h⁻¹ at 50°C and pH 8.5).

There are known from the prior art hydantoin racemases from microorganisms of the genus *Pseudomonas*, *Microbacterium*, *Agrobacterium* and *Arthrobacter* (lit.: JP04271784; EP1188826; Cloning and characterization of genes from *Agrobacterium* sp. IP I-671 involved in hydantoin degradation. Hils, M.; Muench, P.; Altenbuchner, J.; Syldatk, C.; Mattes, R. Applied Microbiology and Biotechnology (2001), 57(5-6), 680-688; A new razemase for 5-monosubstituted hydantoins. Pietzsch, Markus;

Syldatk, Christoph; Wagner, Fritz. *Ann. N. Y. Acad. Sci.* (1992), 672 (*Enzyme Engineering XI*), 478-83. Lickefett, Holger; Krohn, Karsten; Koenig, Wilfried A.; Gehrcke, Barbel; Syldatk, Christoph. *Tetrahedron: Asymmetry* (1993), 4(6), 1129-35; Purification and characterization of the hydantoin racemase of *Pseudomonas* sp. strain NS671 expressed in *Escherichia coli*. Watabe, Ken; Ishikawa, Takahiro; Mukohara, Yukuo; Nakamura, Hiroaki. *J. Bacteriol.* (1992), 174(24), 7989-95).

10 Of the hydantoin racemases from *Arthrobacter aurescens* DSM 3745, *Pseudomonas* sp. NS671 and *Microbacterium liquefaciens*, it is known that these enzymes racemise aliphatic hydantoins, such as, for example, isopropylhydantoin or isobutylhydantoin, only weakly. It is also

15 known that the hydantoin racemases from *Arthrobacter aurescens* DSM 3747 preferentially convert aromatic hydantoins, such as indolylmethylhydantoin or benzylhydantoin, whereas aliphatic hydantoins, such as methylthioethylhydantoin, are converted comparatively weakly or,

20 in the case of isopropylhydantoin, are not converted at all (A new razemase for 5-monosubstituted hydantoins. Pietzsch, Markus; Syldatk, Christoph; Wagner, Fritz. *Ann. N. Y. Acad. Sci.* (1992), 672 (*Enzyme Engineering XI*), 478-83.).

25 The low activity of hydantoin racemases therefore frequently limits the economic potential of this route.

In order to enable as many hydantoin racemases as possible to be checked in a suitable time for their potential to racemise aliphatic hydantoins, the object of the present invention was *inter alia* to provide a suitable screening process for hydantoin racemases. Moreover, the screening process according to the invention should be usable as a component of a mutagenesis process for obtaining new and improved hydantoin racemases. It was also an object of the present invention to provide novel hydantoin racemases

30 which are superior to the hydantoin racemases of the prior

35

art at least in terms of selectivity and/or activity and/or stability.

This object is achieved according to the claims. Claim 1 relates to a screening process for hydantoin racemases.

5 Dependent claims 2 to 4 indicate advantageous embodiments of the screening process. Claim 5 is concerned with a mutagenesis process for the preparation of novel hydantoin racemases using the screening process according to the invention. Claims 6 to 11 relate to novel hydantoin 10 racemases and to the nucleic acid sequences coding therefor and their use. Claims 12 to 14 are directed towards vehicles containing the hydantoin racemases according to the invention, or particular primers for their preparation.

By the provision of a screening process for hydantoin 15 racemases, in which

a) an enantioselective hydantoinase and
b) the hydantoin racemase to be tested, which has a slower conversion rate compared with the hydantoinase under a), are allowed to act on

20 c) a chiral hydantoin, which is used in the opposite enantiomERICALLY enriched form to the selectivity of the hydantoinase, and
d) the resulting N-carbamoyl-amino acid or the freed protons are detected in a time-dependent manner,

25 it becomes possible in a surprisingly simple and yet advantageous manner to check a large number of hydantoin racemases in a short time for their ability to racemise hydantoins in an improved manner.

By the use of an L-enantiomer of a 5'-monosubstituted 30 hydantoin and the use of a D-selective hydantoinase which, on the basis of its enantioselectivity, preferably rapidly hydrolyses the resulting D-enantiomer of the hydantoin, the racemisation rate and hence the activity of the hydantoin racemase can be measured in a simple manner by the 35 formation of the N-carbamoyl-D-amino acid or by freed

protons. The N-carbamoyl-amino acid can be quantified by methods known to the person skilled in the art, such as, for example, HPLC or colorimetric methods. Quantification via protons can be carried out in a simple manner via pH indicators, preferably cresol red. It should be noted that both D- and L-enantiomers of hydantoins having different optionally aliphatic 5'-substituents can be used in the process. When D-hydantoins are used, corresponding L-selective hydantoinases are to be used in the screening process.

In the process according to the invention there are advantageously used aliphatic hydantoins substituted in the 5'-position. In this context, aliphatically substituted hydantoins are understood to mean a system which has in the 15 5'-position on the hydantoin heterocycle a radical which is bonded to the heterocycle via a carbon atom having sp^3 - hybridisation. Preferred 5'-substituents are methyl, ethyl, butyl, propyl, tertiary butyl, isopropyl and isobutyl. Ethylhydantoin is very particularly preferred.

20 There may be used as hydantoinases any hydantoinases known in the literature which enantioselectively hydrolyse the hydantoin enantiomer formed via the hydantoin racemase, it being necessary for this hydrolysis to be more rapid than the racemisation rate. Preferred hydantoinases are the 25 commercial hydantoinases 1 & 2 from Roche, the hydantoinases of the genera *Agrobacterium*, *Arthrobacter*, *Bacillus*, *Pseudomonas*, *Flavobacterium*, *Pasteurella*, *Microbacterium*, *Vigna*, *Ochrobactrum*, *Methanococcus*, *Burkholderia* and *Streptomyces*. (Hils, M.; Muench, P.; 30 Altenbuchner, J.; Syldatk, C.; Mattes, R. Cloning and characterization of genes from *Agrobacterium* sp. IP I-671 involved in hydantoin degradation. *Applied Microbiology and Biotechnology* (2001), 57(5-6), 680-688. Soong, C.-L.; Ogawa, J.; Shimizu, S. Cyclic ureide and imide metabolism 35 in microorganisms producing a D-hydantoinase useful for D-

amino acid production. *Journal of Molecular Catalysis B: Enzymatic* (2001), 12(1-6), 61-70. Wiese, Anja; Wilms, Burkhard; Syldatk, Christoph; Mattes, Ralf; Altenbuchner, Josef. Cloning, nucleotide sequence and expression of a 5 hydantoinase and carbamoylase gene from *Arthrobacter aurescens* DSM 3745 in *Escherichia coli* and comparison with the corresponding genes from *Arthrobacter aurescens* DSM 3747. *Applied Microbiology and Biotechnology* (2001), 55(6), 750-757. Yin, Bang-Ding; Chen, Yi-Chuan; Lin, Sung- 10 Chyr; Hsu, Wen-Hwei. Production of D-amino acid precursors with permeabilized recombinant *Escherichia coli* with D-hydantoinase activity. *Process Biochemistry* (Oxford) (2000), 35(9), 915-921. Park, Joo-Ho; Kim, Geun-Joong; Lee, Seung-Goo; Lee, Dong-Cheol; Kim, Hak-Sung. 15 Purification and characterization of thermostable D-hydantoinase from *Bacillus thermocatenulatus* GH-2. *Applied Biochemistry and Biotechnology* (1999), 81(1), 53-65; Pozo, C.; Rodelas, B.; de la Escalera, S.; Gonzalez-Lopez, J. D,L-Hydantoinase activity of an *Ochrobactrum anthropi* 20 strain. *Journal of Applied Microbiology* (2002), 92(6), 1028-1034; Chung, Ji Hyung; Back, Jung Ho; Lim, Jae-Hwan; Park, Young In; Han, Ye Sun. Thermostable hydantoinase from a hyperthermophilic archaeon, *Methanococcus jannaschii*. *Enzyme and Microbial Technology* (2002), 25 30(7), 867-874; Xu, Zhen; Jiang, Weihong; Jiao, Ruishen; Yang, Yunliu. Cloning, sequencing and high expression in *Escherichia coli* of D-hydantoinase gene from *Burkholderia pickettii*. *Shengwu Gongcheng Xuebao* (2002), 18(2), 149-154; Las Heras-Vazquez, Francisco Javier; Martinez- 30 Rodriguez, Sergio; Mingorance-Cazorla, Lydia; Clemente-Jimenez, Josefa Maria; Rodriguez-Vico, Felipe. Overexpression and characterization of hydantoin racemase from *Agrobacterium tumefaciens* C58. *Biochemical and Biophysical Research Communications* (2003), 303(2), 35 541-547; DE 3535987; EP 1275723; US 6087136; WO 0281626; US 2002045238; DE 4328829; WO 9400577; WO 9321336; JP

04325093; NL 9001680; JP 2003024074; WO 0272841; WO 0119982; WO 9620275).

The use of the hydantoinase from *Arthrobacter crystallopoietes*, especially from DSM 20117, is very
5 particularly preferred.

As already indicated, the conversion rate of the hydantoinase should be superior to that of the racemase. The ratio of the rate constants of the hydantoinase to the hydantoin racemase (k_{hyd}/k_{rac}) is preferably > 2 ,
10 particularly preferably > 10 and very particularly preferably > 50 .

The invention also provides a process for the preparation of improved hydantoin racemases, which is distinguished by the fact that

- 15 a) the nucleic acid sequence coding for the hydantoin racemase is subjected to a mutagenesis,
- b) the nucleic acid sequences obtainable from a) are cloned into a suitable vector and the vector is transferred into a suitable expression system, and
- 20 c) the resulting hydantoin racemases having improved activity and/or selectivity and/or stability are detected by means of a screening process according to the invention and isolated.

There may be used as starting genes for the mutagenesis of the hydantoin racemases any known hydantoin racemase genes mentioned in the listed literature. Preference is given to the hydantoin racemase genes of *Arthrobacter*, *Pseudomonas*, *Agrobacterium* and *Micrococcus* (Wiese A; Pietzsch M; Syldatk C; Mattes R; Altenbuchner J Hydantoin racemase from 30 *Arthrobacter aurescens* DSM 3747: heterologous expression, purification and characterization. JOURNAL OF BIOTECHNOLOGY (2000 Jul 14), 80(3), 217-30; Watabe K; Ishikawa T; Mukohara Y; Nakamura H Purification and characterization of the hydantoin racemase of *Pseudomonas* sp. strain NS671 expressed in *Escherichia coli*. JOURNAL OF

BACTERIOLOGY (1992 Dec), 174(24), 7989-95; Las Heras-Vazquez, Francisco Javier; Martinez-Rodriguez, Sergio; Mingorance-Cazorla, Lydia; Clemente-Jimenez, Josefa Maria; Rodriguez-Vico, Felipe. Overexpression and

5 characterization of hydantoin racemase from *Agrobacterium tumefaciens* C58. Biochemical and Biophysical Research Communications (2003), 303(2), 541-547; EP 1188826). Very particular preference is given to the hydantoin racemase gene from *Arthrobacter aurescens*, which codes for 10 the protein sequence in Seq.ID.No. 2.

For the mutagenesis of the hydantoin racemase there may be used any methods known in the literature, such as, for example, random mutagenesis, saturation mutagenesis, cassette mutagenesis or recombination methods (May, Oliver; 15 Voigt, Christopher A.; Arnold, Frances H. Enzyme engineering by directed evolution. Enzyme Catalysis in Organic Synthesis (2nd Edition) (2002), 1 95-138; Bio/Technology 1991, 9, 1073-1077; Horwitz, M. and Loeb, L., Promoters Selected From Random DNA-Sequences, Proc Natl Acad Sci USA 83, 1986, 7405-7409; Dube, D. and L. Loeb, Mutants Generated By The Insertion Of Random 20 Oligonucleotides Into The Active-Site Of The Beta-Lactamase Gene, Biochemistry 1989, 28, 5703-5707; Stemmer, P.C., Rapid evolution of a protein in vitro by DNA shuffling, 25 Nature 1994, 370, 389-391 and Stemmer, P.C., DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution. Proc Natl Acad Sci USA 91, 1994, 10747-10751).

The cloning and expression can be carried out as in the 30 literature mentioned hereinbelow. The process can be carried out several times in succession, optionally with varying mutagenesis strategies.

The invention also provides rec-polypeptides or the nucleic acid sequences coding therefor, which are obtainable by the 35 mutagenesis process mentioned above.

Another aspect of the invention is the use of the polypeptides so prepared in the preparation of chiral enantiomerically enriched N-carbamoyl-amino acids or amino acids. The nucleic acid sequences prepared according to the 5 invention can be used in the preparation of whole cell catalysts.

Hydantoin racemases which have in position 79 an amino acid substitution with an amino acid selected from the group consisting of A, R, N, D, C, Q, E, H, I, L, K, M, F, P, S, 10 T, Y and V also form part of the present invention. It is interesting that the amino acids surrounding this position are retained completely for many hydantoin racemases. The consensus sequence reads: FX₁DX₂GL (Seq. ID. No. 1), wherein X₂ represents P or T and X₁ represents W or G. Preferred 15 mutants therefore contain the above-mentioned consensus sequence, X₁ preferably representing an amino acid selected from the group consisting of A, R, N, D, C, Q, E, H, I, L, K, M, F, P, S, T, Y and V. X₁ corresponds to position 79. Preferred mutants are shown in Table 2.

20 Table 2:

Mutant name	Mutation (codon)	Mutation X ₁ (amino acid)	Activity change	Seq. ID. No.
3CH11	GGG -> GAG	G79E	2	5
1BG7	GGG -> AGG	G79R	2	3
BB5	GGG -> TTG	G79L	4	9
AE3	GGG -> CAG	G79Q	4	7

Further extremely advantageous combinations of X₁ and X₂ hydantoin racemases are listed in Table 3 below.

Table 3: Advantageous combinations of X_1 and X_2 in the consensus motif FX_1DX_2GL

X_1	L	E	Q	R	X_1	L	E	Q	R
X_2	P	P	P	P		T	T	T	T

It is particularly advantageous if the hydantoin racemases
 5 contain the above-mentioned consensus region and
 additionally exhibit a homology of >40% with the hydantoin
 racemase from DSM 20117.

The invention also provides isolated nucleic acid sequences
 coding for a hydantoin racemase selected from the group:

- 10 a) a nucleic acid sequence coding for a hydantoin racemase according to the invention,
- b) a nucleic acid sequence which hybridises under stringent conditions with the nucleic acid sequence coding for a hydantoin racemase according to the invention or with
 15 the sequence complementary thereto,
- c) a nucleic acid sequence according to Seq.ID.No. 3, 5, 7 or 9 or a nucleic acid sequence having a homology of > 80% therewith,
- d) a nucleic acid sequence containing 15 successive
 20 nucleotides of sequences Seq.ID.No. 3, 5, 7 or 9.

With regard to point d), it is preferred for the nucleotide sequence according to the invention to contain 20, more preferably 25, yet more preferably 30, 31, 32, 33, 34 and most preferably more than 34 identical consecutive nucleic
 25 acids of the sequences Seq.ID.No. 3, 5, 7 or 9.

As mentioned, the invention also includes nucleic acid sequences which hybridise under stringent conditions with the single-strand nucleic acid sequences according to the invention or with their complementary single-strand nucleic
 30 acid sequences (b), or nucleic acid sequences which are alike in sequence sections (d). Particular gene probes or

the primers necessary for a PCR, for example, are to be regarded as such.

Coupling of hydantoin racemase and hydantoinase and optionally carbamoylase can be carried out by bringing together the free or immobilised enzymes. However, it is preferred for the hydantoinase to be expressed in the same cell together with the hydantoin racemase and/or the carbamoylase (whole cell catalyst).

The nucleic acid sequences according to the invention can therefore be cloned into a whole cell catalyst as a constituent of a gene in a manner analogous to that in DE10234764 and the literature cited therein.

Provided that the latter then also contains genes for a hydantoinase and/or carbamoylase, it is capable of converting racemic hydantoins completely into enantiomerically enriched amino acids. Without a cloned carbamoylase gene, the reaction stops at the stage of the N-carbamoyl-amino acids.

The host organism used is preferably an organism as mentioned in DE10155928. The advantage of such an organism is the simultaneous expression of all the enzymes involved, with which only a rec-organism must be used for the total reaction.

In order to match the expression of the enzymes in respect of their conversion rates, the corresponding coding nucleic acid sequences can be cloned into different plasmids with different copy numbers and/or promoters of different strengths can be used for a different strength of expression of the nucleic acid sequences. In such matched enzyme systems, there is advantageously no accumulation of an intermediate compound which may have an inhibiting action, and the reaction under consideration can proceed at an optimum overall rate. This is sufficiently well known to the person skilled in the art, however (Gellissen, G.; Piontek, M.; Dahlems, U.; Jenzelewski, V.; Gavagan, J. W.; DiCosimo, R.; Anton, D. L.; Janowicz, Z. A. (1996),

Recombinant *Hansenula polymorpha* as a biocatalyst.

Coexpression of the spinach glycolate oxidase (GO) and the *S. cerevisiae* catalase T (CTT1) gene, *Appl. Microbiol.*

Biotechnol. 46, 46-54; Farwick, M.; London, M.; Dohmen, J.;

5 Dahlems, U.; Gellissen, G.; Strasser, A. W.; DE19920712).

The preparation of such a whole cell catalyst is sufficiently well known to the person skilled in the art (Sambrook, J.; Fritsch, E. F. and Maniatis, T. (1989), *Molecular cloning: a laboratory manual*, 2nd ed., Cold

10 Spring Harbor Laboratory Press, New York; Balbas, P. and Bolivar, F. (1990), *Design and construction of expression plasmid vectors in E. coli*, *Methods Enzymol.* 185, 14-37;

Rodriguez, R.L. and Denhardt, D. T (eds) (1988), *Vectors: a survey of molecular cloning vectors and their uses*,

15 205-225, Butterworth, Stoneham).

In a next embodiment, the invention relates to plasmids or vectors containing one or more of the nucleic acid sequences according to the invention.

Suitable plasmids or vectors are in principle any forms

20 available to the person skilled in the art for this purpose. Such plasmids and vectors will be found, for example, in Studier et al. (Studier, W. F.; Rosenberg

A. H.; Dunn J. J.; Dubendorff J. W.; (1990), *Use of the T7 RNA polymerase to direct expression of cloned genes*,

25 *Methods Enzymol.* 185, 61-89) or the brochures of Novagen, Promega, New England Biolabs, Clontech or Gibco BRL.

Further preferred plasmids and vectors can be found in:

Glover, D. M. (1985), *DNA cloning: a practical approach*, Vol. I-III, IRL Press Ltd., Oxford; Rodriguez, R.L. and

30 Denhardt, D. T (eds) (1988), *Vectors: a survey of molecular cloning vectors and their uses*, 179-204, Butterworth, Stoneham; Goeddel, D. V. (1990), *Systems for heterologous gene expression*, *Methods Enzymol.* 185, 3-7; Sambrook, J.;

Fritsch, E. F. and Maniatis, T. (1989), *Molecular cloning: a laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory

35 Press, New York.

Plasmids with which the gene construct containing the nucleic acid according to the invention can be cloned into the host organism in a very preferred manner are derivatives of pUC18 and pUC19 (Roche Biochemicals),

5 pKK-177-3H (Roche Biochemicals), pBTac2 (Roche Biochemicals), pKK223-3 (Amersham Pharmacia Biotech), pKK-233-3 (Stratagene) or PET (Novagen). Further preferred plasmids are pBR322 (DSM3879), pACYC184 (DSM4439) and pSC101 (DSM6202), which can be obtained from DSMZ-Deutsche 10 Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany.

The invention is likewise directed towards microorganisms containing one or more nucleic acid sequences according to the invention. The microorganism into which the plasmids

15 containing the nucleic acid sequences according to the invention are cloned serves to multiply and obtain a sufficient amount of the recombinant enzyme. The processes therefor are well known to the person skilled in the art (Sambrook, J.; Fritsch, E. F. and Maniatis, T. (1989),

20 Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York). In principle, there can be used as microorganisms any organisms suitable to the person skilled in the art for this purpose, such as, for example, yeasts, such as *Hansenula polymorpha*, *Pichia* sp., *Saccharomyces cerevisiae*, prokaryotes, such as *E. coli*, *Bacillus subtilis*, or eukaryotes, such as mammalian 25 cells, insect cells. *E. coli* strains are preferably to be used for this purpose. Very particular preference is given to: *E. coli* XL1 Blue, W3110, DSM14459 (PCT/US00/08159), NM 30 522, JM101, JM109, JM105, RR1, DH5 α , TOP 10 γ and HB101.

Plasmids with which the gene construct containing the nucleic acid according to the invention is preferably cloned into the host organism are indicated above.

A following aspect of the invention is directed towards 35 primers for the preparation of the gene sequences according to the invention by means of any type of PCR. Included are

the sense and antisense primers coding for the corresponding amino acid sequences, or complementary DNA sequences. Suitable primers can in principle be obtained by processes known to the person skilled in the art. The

5 location of the primers according to the invention is carried out by comparison with known DNA sequences or by translation of the amino acid sequences under consideration into the preferred codon of the organism in question (e.g. for *Streptomyces*: Wright F. and Bibb M. J. (1992), Codon 10 usage in the G+C-rich *Streptomyces* genome, Gene 113, 55-65). Similarities in the amino acid sequence of proteins of so-called superfamilies are likewise of use therefor (Firestone, S. M.; Nixon, A. E.; Benkovic, S. J. (1996), Threading your way to protein function, *Chem. Biol.* 3, 15 779-783). Further information hereon can be found in Gait, M. J. (1984), *Oligonucleotide synthesis: a practical approach*, IRL Press Ltd., Oxford; Innis, M. A.; Gelfound, D. H.; Sninsky, J. J. and White, T.J. (1990), *PCR Protocols: A guide to methods and applications*, Academic 20 Press Inc., San Diego.

Preferred primers are those of Seq.ID.No. 11 and 12.

As already indicated, the enzymes under consideration (hydantoin racemase, hydantoinases and/or carbamoylases) can be used in free form as homogeneously purified

25 compounds or as an enzyme prepared by recombinant methods (rec-). The enzymes may also be used as a constituent of an intact guest organism or in conjunction with the cell mass of the host organism which has been opened up and highly purified as desired.

30 It is also possible to use the enzymes in immobilised form (Sharma B. P.; Bailey L. F. and Messing R. A. (1982), *Immobilisierte Biomaterialien - Techniken und Anwendungen*, Angew. Chem. 94, 836-852). Immobilisation is preferably carried out by lyophilisation (Paradkar, V. M.; Dordick, J. 35 S. (1994), *Aqueous-Like Activity of α -Chymotrypsin Dissolved in Nearly Anhydrous Organic Solvents*, J. Am.

Chem. Soc. 116, 5009-5010; Mori, T.; Okahata, Y. (1997), A variety of lipi-coated glycoside hydrolases as effective glycosyl transfer catalysts in homogeneous organic solvents, Tetrahedron Lett. 38, 1971-1974; Otamiri, M.; 5 Adlercreutz, P.; Matthiasson, B. (1992), Complex formation between chymotrypsin and ethyl cellulose as a means to solbilize the enzyme in active form in toluene, Biocatalysis 6, 291-305). Very special preference is given to lyophilisation in the presence of surface-active 10 substances, such as Aerosol OT or polyvinylpyrrolidone or polyethylene glycol (PEG) or Brij 52 (diethylene glycol monocetyl ether) (Kamiya, N.; Okazaki, S.-Y.; Goto, M. (1997), Surfactant-horseradish peroxidase complex catalytically active in anhydrous benzene, Biotechnol. 15 Tech. 11, 375-378). Very special preference is given to immobilisation on Eupergit®, especially Eupergit C® and Eupergit 250L® (Röhm) (Eupergit.RTM. C, a carrier for immobilization of enzymes of industrial potential. Katchalski-Katzir, E.; Kraemer, D. 20 M. Journal of Molecular Catalysis B: Enzymatic (2000), 10(1-3), 157-176.). Also preferred is immobilisation on Ni-NTA in combination with the polypeptide supplemented with the His tag (hexahistidine) (Purification of proteins using polyhistidine 25 affinity tags. Bornhorst, Joshua A.; Falke, Joseph J. Methods in Enzymology (2000), 326, 245-254). Use as CLECs is also conceivable (St. Clair, N.; Wang, Y.-F.; Margolin, A. L. (2000), Cofactor-bound cross-linked enzyme crystals (CLEC) of alcohol dehydrogenase, Angew. Chem. Int. Ed. 39, 30 380-383). By means of these measures it can be possible to generate from polypeptides that are rendered unstable by organic solvents polypeptides that are stable and can work in mixtures of aqueous and organic solvents or in wholly 35 organic solvents.

Whole cell catalysts are generally used in the form of free or immobilised cells. For this purpose, the active cell mass is re-suspended in a hydantoin-containing solution. The cell concentration is from 1 to 100 g/l. The 5 concentration of hydantoin is from 0.1 to 2 molar. H₂O is preferably used as the solvent, but mixtures of organic solvents and H₂O can also be used. The pH value is either not controlled or is maintained between pH 6 and pH 10 by means of conventional buffers or by continuous pH 10 monitoring. The reaction temperature is typically from 20°C to 90°C. In dependence on the hydantoinase used, divalent metal ions are added in concentrations of from 0.1 to 5 mM. Preferred metal ions are Mn²⁺, Zn²⁺ or Co²⁺. With regard to the use of the individual enzymes, an 15 equivalent procedure can be employed.

The products prepared by the use of the hydantoin racemases according to the invention in the manner as described, for example, above are worked up by conventional methods. However, working up by ion exchange chromatography is 20 preferred. As a result, the product is freed of the salts formed in the reaction. The eluate is optionally clarified using activated carbon and the resulting enantiomerically enriched amino acid or N-carbamoyl-amino acid is precipitated by concentration of the solvent and dried. 25 Coupling of an enzymatic racemisation with an enantioselective hydrolysis for the screening of hydantoin racemase activities has not hitherto been used to produce improved hydantoin racemases. For the process according to the invention to be applied particularly successfully, 30 several requirements should be met:

1. The chemical racemisation rate of the enantiomerically pure hydantoin used in the screening must be very much lower than the rate of the enzymatically catalysed reaction.

2. The enantioselective hydrolysis by means of the hydantoinase must take place very much more rapidly than the enzymatic racemisation of the hydantoin.

For aliphatically substituted hydantoins, point 1 is met 5 owing to their slow chemical racemisation. Point 2 can be fulfilled by a targeted selection of suitable hydantoinases (see hereinabove).

10 The present invention is not rendered obvious by the statements made in the prior art, because no indications are to be found therein relating to the requirements mentioned hereinbefore.

15 All the indicated mutants exhibit a mutation at amino acid position 79, which for the first time indicates the importance of this position for the enzyme function. It is interesting that the amino acids surrounding this position are retained completely for all known hydantoin racemases. This shows that, for other hydantoin racemases which contain the above-described sequence motif and exhibit a high degree of homology (>40% sequence identity), improved 20 enzyme variants can be produced by site-specific mutagenesis at position 79, which could not hitherto be derived from the prior art.

25 Within the scope of the invention, the expression optically enriched (enantiomerically enriched) compounds is understood to mean the presence of one optical antipode in admixture with the other in >50 mol.%.

The expression nucleic acid sequences includes all types of single-strand or double-strand DNA as well as RNA or mixtures thereof.

30 According to the invention, the improvement in the activity and/or selectivity and/or stability means that the polypeptides are more active and/or more selective or less selective or more stable under the reaction conditions.

While the activity and the stability of the enzymes should naturally be as high as possible for technical application, the selectivity is said to be improved when either the substrate selectivity falls but the enantioselectivity of 5 the enzymes is increased.

According to the invention, the claimed polypeptides and the nucleic acid sequences also include those sequences which exhibit a homology (excluding natural degeneration) of greater than 70% (in respect of the nucleic acid 10 sequence) or > 40% or 80% (in respect of the polypeptides), preferably greater than 90%, 91%, 92%, 93% or 94%, more preferably greater than 95% or 96% and particularly preferably greater than 97%, 98% or 99%, with one of these 15 sequences, provided that the mode of action or purpose of such a sequence is retained. The expression "homology" (or identity) as used herein can be defined by the equation $H (\%) = [1 - V/X] \times 100$, where H means homology, X is the total number of nucleobases/amino acids in the comparison sequence and V is the number of different nucleobases/amino 20 acids of the sequence under consideration relative to the comparison sequence. In any case, the expression nucleic acid sequences coding for polypeptides includes all sequences that appear possible according to the degeneration of the genetic code.

25 The expression "under stringent conditions" is understood herein as described in Sambrook *et al.* (Sambrook, J.; Fritsch, E. F. and Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York). A stringent hybridisation according to 30 the present invention is preferably present when, after washing for one hour with 1 x SSC (150 mM sodium chloride, 15 mM sodium citrate, pH 7.0) and 0.1 % SDS (sodium dodecylsulfate) at 50°C, preferably at 55°C, more preferably at 62°C and most preferably at 68°C, and more preferably for 35 1 hour with 0.2 x SSC and 0.1 % SDS at 50°C, more preferably

at 55°C, yet more preferably at 62°C and most preferably at 68°C, a positive hybridisation signal is still observed.

The literature references cited in this specification are incorporated in the disclosure by reference.

- 5 The organism *Arthrobacter aurescens* DSM3747 was deposited with Deutsche Sammlung für Mikroorganismen GmbH, Mascheroder Weg 1b, 38124 Braunschweig by Rütgerswerke Aktiengesellschaft on 28.05.86.

Examples**Example 1: Production of hydantoin racemase mutants - random mutagenesis**

5 0.25 ng of the vector pOM21 (plasmid map see Fig.1; sequence see Seq.ID.No.13) (PCT/US00/08159) was used as template in a 100 µl PCR reaction mix consisting of PCR buffer (10 mM Tris, 1.5 mM MgCl₂, 50 mM KCl, pH 8.5), 200 µM dTTP, 200 µM dGTP, 200 µM dATP, 200 µM dCTP,

10 50 pmol. of the respective primer (see Seq.ID.No.11 and 12) and 2.5 U Taq polymerase (Roche). After 30 cycles, the amplified product was purified by means of gel extraction (QiaexII gel extraction kit) and subcloned into the vector pOM21 by means of the restriction enzymes NdeI and PstI.

15 The ligation product was used for the transformation of hydantoinase-positive strains (see Example 2).

Example 2: Preparation of hydantoinase-positive strains and of a mutant library

Chemically competent *E. coli* JM109 (e.g. from Promega) were transformed with 10 ng of the plasmid pDHYD (see Fig.2; see Seq.ID.No. 15), which carries the D-hydantoinase gene from *Arthrobacter crystallopoietes* DSM20117 under the control of a rhamnose promoter. The complete sequence of the plasmid is shown in Seq.ID.No. 15. The hydantoinase-positive strain so produced was in turn rendered chemically competent and transformed for the preparation of the mutant library with the ligation product of the hydantoin racemase random mutagenesis from Example 1. The colonies of the mutant library were spread onto ampicillin- and chloramphenicol-containing agar plates and then subjected to a screening, which is described in Example 3.

Example 3: Screening for hydantoin racemase mutants having improved enzyme properties

Individual colonies of the mutant library were inoculated in 96-well plates which were filled with 100 μ l per well of 5 LB medium (5 g/l yeast extract, 10 g/l tryptone, 10 g/l NaCl) supplemented with rhamnose (2 g/l) and ZnCl₂ (1 mM). The plates were incubated for 20 hours at 30°C. 100 μ l of screening substrate (100 mM L-ethylhydantoin, 50 mg/l cresol red, pH 8.5) were then added to each well and the 10 plates were incubated for 4 hours at 20°C. Wells having improved hydantoin racemase mutants could be identified directly with the eye by means of a more intense yellow colouration compared with the wild type, or using a spectral photometer at 580 nm.

15 Example 4: Characterisation of hydantoin racemase mutants having improved enzyme properties

The racemase mutants identified in the screening were subsequently tested by means of HPLC analysis for their activity in comparison with the wild type, and the 20 corresponding mutations were determined by means of sequencing. For this purpose, plasmids were isolated from individual colonies of the different clones (Qiagen Mini-Prep Kit) and sequenced. The same clones were used to produce active biomass. An overnight culture (OD₆₀₀=4) of 25 the respective clones was to this end diluted 1:100 in 100 ml of LB medium (5 g/l yeast extract, 10 g/l tryptone, 10 g/l NaCl) supplemented with rhamnose (2 g/l) and ZnCl₂ (1 mM) and incubated for 18 hours at 30°C and 250 rpm. The biomass was pelletised by centrifugation (10 min, 10,000 g) 30 and the supernatant was discarded. 2 g of active biomass were then re-suspended in 50 ml of the substrate solution (100 mM L-ethylhydantoin, pH 8.5) and incubated at 37°C. Samples were taken after various times, the biomass was separated off by centrifugation (5 min, 13,000 rpm) and the

supernatant was analysed by means of HPLC for the concentration of the N-carbamoyl-aminobutyric acid formed.

Example 5: Preparation of L-amino acids using improved hydantoin racemases

5 A strain of *E. coli* JM109 transformed with pOM21-BB5 and pOM22 Fig. 3 (see Seq.ID.No.14) (PCT/US00/08159) was incubated at 30°C for 18 hours, with shaking (250 rpm), in LB medium which contained ampicillin (100 µg/l) and chloramphenicol (50 µg/l) and to which 2 g/l of rhamnose
10 had been added. The biomass was pelletised by centrifugation and re-suspended in a corresponding volume of 100 mM DL-ethylhydantoin solution, pH 8.5, and 1 mM CoCl₂, so that a cell concentration of 30 g/l was obtained. This reaction solution was incubated for 10 hours at 37°C.
15 The cells were then separated off by centrifugation (30 min, 5000 g) and the clear supernatant was analysed by means of HPLC for the resulting amino acid. For working up the resulting amino acid, the volume of the supernatant was reduced to half, and methanol was added 1:2. The
20 precipitated amino acid was then filtered off and dried. The total yield of the amino acid was >60%.

Example 6: Preparation of D-amino acids using improved hydantoin racemases

A strain of *E. coli* JM109 transformed with pOM21-BB5 and pJAVIER16 Fig. 4 (see Seq.ID.No.16) was incubated at 30°C for 18 hours, with shaking (250 rpm), in LB medium which contained ampicillin (100 µg/l) and chloramphenicol (50 µg/l) and to which 2 g/l of rhamnose had been added. The biomass was pelletised by centrifugation and re-suspended
25 in a corresponding volume of 100 mM DL-ethylhydantoin solution, pH 8.5, and 1 mM CoCl₂, so that a cell concentration of 30 g/l was obtained. This reaction solution was incubated for 10 hours at 37°C. The cells were then separated off by centrifugation (30 min, 5000 g) and

the clear supernatant was analysed by means of HPLC for the resulting amino acid. For working up the resulting amino acid, the volume of the supernatant was reduced to half, and methanol was added 1:2. The precipitated amino acid was 5 then filtered off and dried. The total yield of the amino acid was >60%.

Patent claims:

1. Screening process for hydantoin racemases, characterised in that
 - a) an enantioselective hydantoinase and
 - 5 b) the hydantoin racemase to be tested, which has a slower conversion rate compared with the hydantoinase under a), are allowed to act on
 - c) a chiral hydantoin, which is used in the opposite enantiomerically enriched form to the selectivity 10 of the hydantoinase, and
 - d) the resulting N-carbamoyl-amino acid or the freed protons are detected in a time-dependent manner.
2. Process according to claim 1, characterised in that 15 an aliphatically substituted hydantoin is used.
3. Process according to one or more of the preceding claims, characterised in that a hydantoinase from *Arthrobacter crystallopoietes* is 20 used.
4. Process according to one or more of the preceding claims, characterised in that the ratio of the rate constants of the hydantoinase to 25 the hydantoin racemase (k_{hyd}/k_{rac}) is > 2 .
5. Process for the preparation of improved hydantoin racemases, characterised in that
 - a) the nucleic acid sequence coding for the hydantoin racemase is subjected to a mutagenesis,
 - 30 b) the nucleic acid sequences obtainable from a) are cloned into a suitable vector and the vector is transferred into a suitable expression system, and

c) the resulting hydantoin racemases having improved activity and/or selectivity and/or stability are detected by means of a process according to one or more of claims 1 to 4 and isolated.

- 5 6. rec-Polypeptides or nucleic acid sequences coding therefor obtainable according to claim 5.
7. Use of the polypeptides according to claim 6 in the preparation of enantiomerically enriched N-carbamoyl-amino acid or amino acids.
- 10 8. Use of the nucleic acid sequences according to claim 6 in the preparation of whole cell catalysts.
9. Hydantoin racemase containing in position 79 an amino acid substitution with an amino acid selected from the group consisting of A, R, N, D, C, Q, E, H, I, L, K, M, F, P, S, T, Y and V.
- 15 10. Hydantoin racemases containing the consensus sequence FX₁DX₂GL (Seq. 1), wherein X₂ represents P or T and X₁ represents in position 79 an amino acid selected from the group A, R, N, D, C, Q, E, H, I, L, K, M, F, P, S, T, Y or V.
- 20 11. Isolated nucleic acid sequence coding for a hydantoin racemase selected from the group:
 - a) a nucleic acid sequence coding for a hydantoin racemase according to claim 9 and/or 10,
 - 25 b) a nucleic acid sequence which hybridises under stringent conditions with the nucleic acid sequence coding for a hydantoin racemase according to claim 9 and/or 10 or with the sequence complementary thereto,
 - 30 c) a nucleic acid sequence according to Seq.ID.No. 3, 5, 7 or 9 or a nucleic acid sequence having a homology of > 80% therewith,
 - d) a nucleic acid sequence containing 15 successive

nucleotides of the sequences Seq.ID.No. 3, 5, 7 or 9.

12. Whole cell catalyst containing a cloned gene for a hydantoin racemase according to claims 9 and/or 10.
- 5 13. Plasmids, vectors or microorganisms containing a nucleic acid sequence according to claim 9 and/or 10.
14. Primers for the preparation of the nucleic acid sequences according to claim 9 and/or 10 by means of PCR.

Fig. 1:

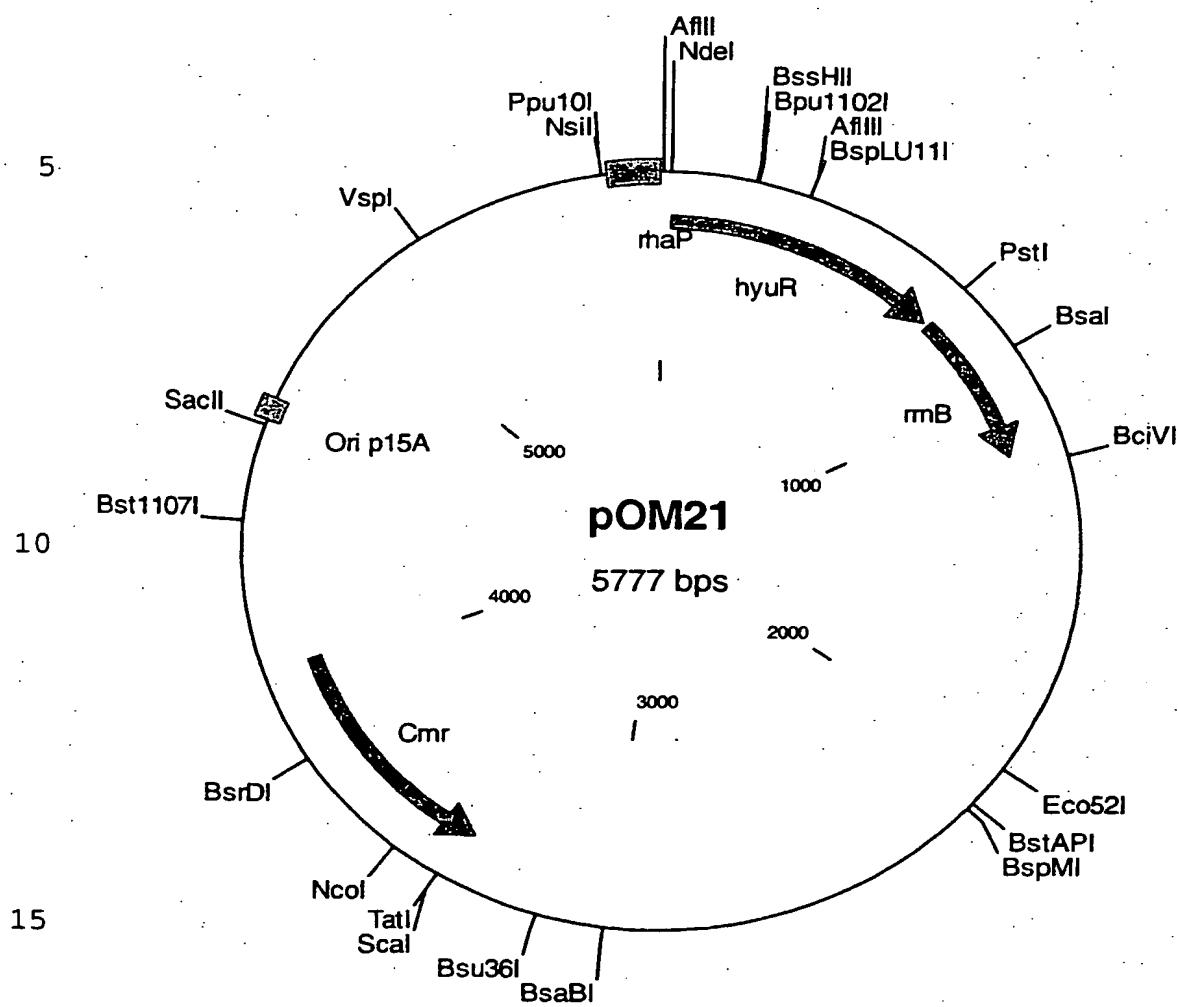


Fig. 2:

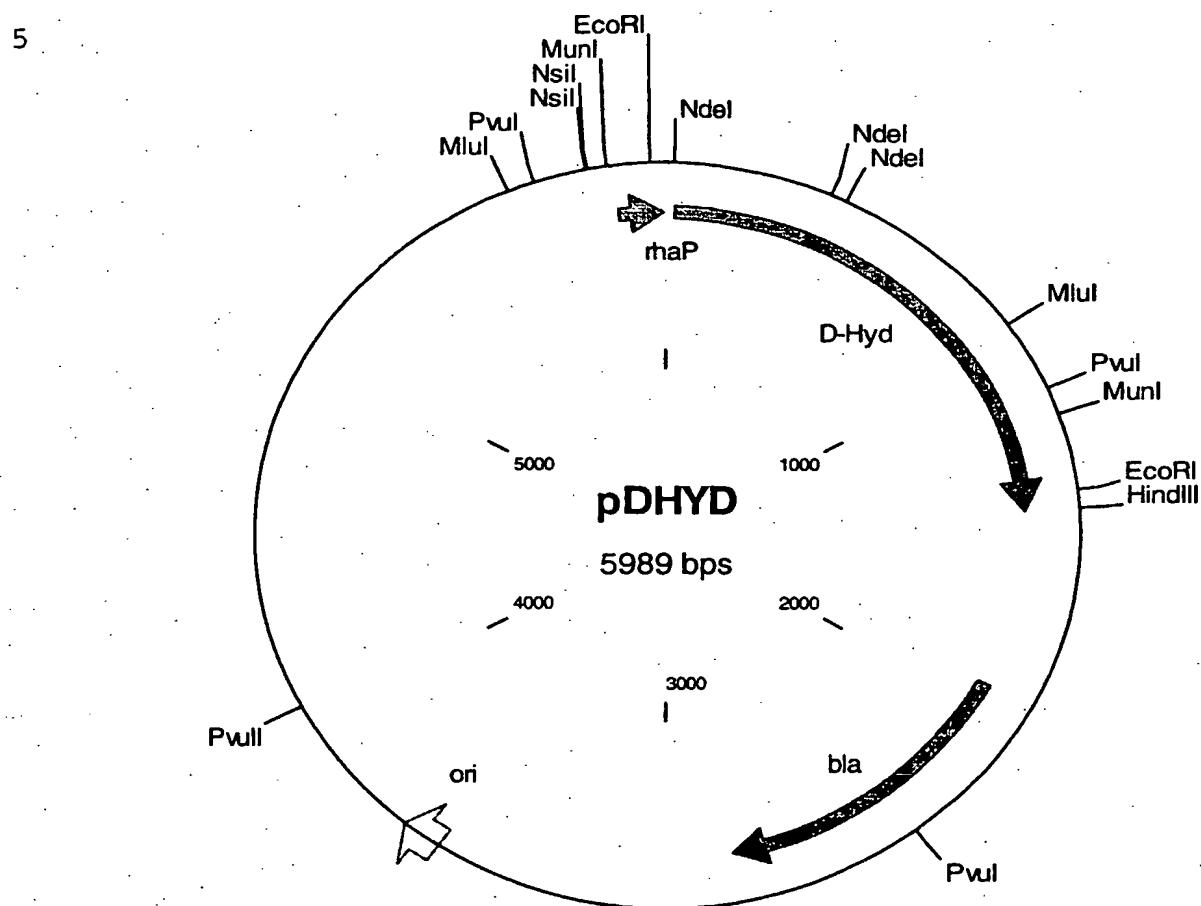


Fig: 3

5

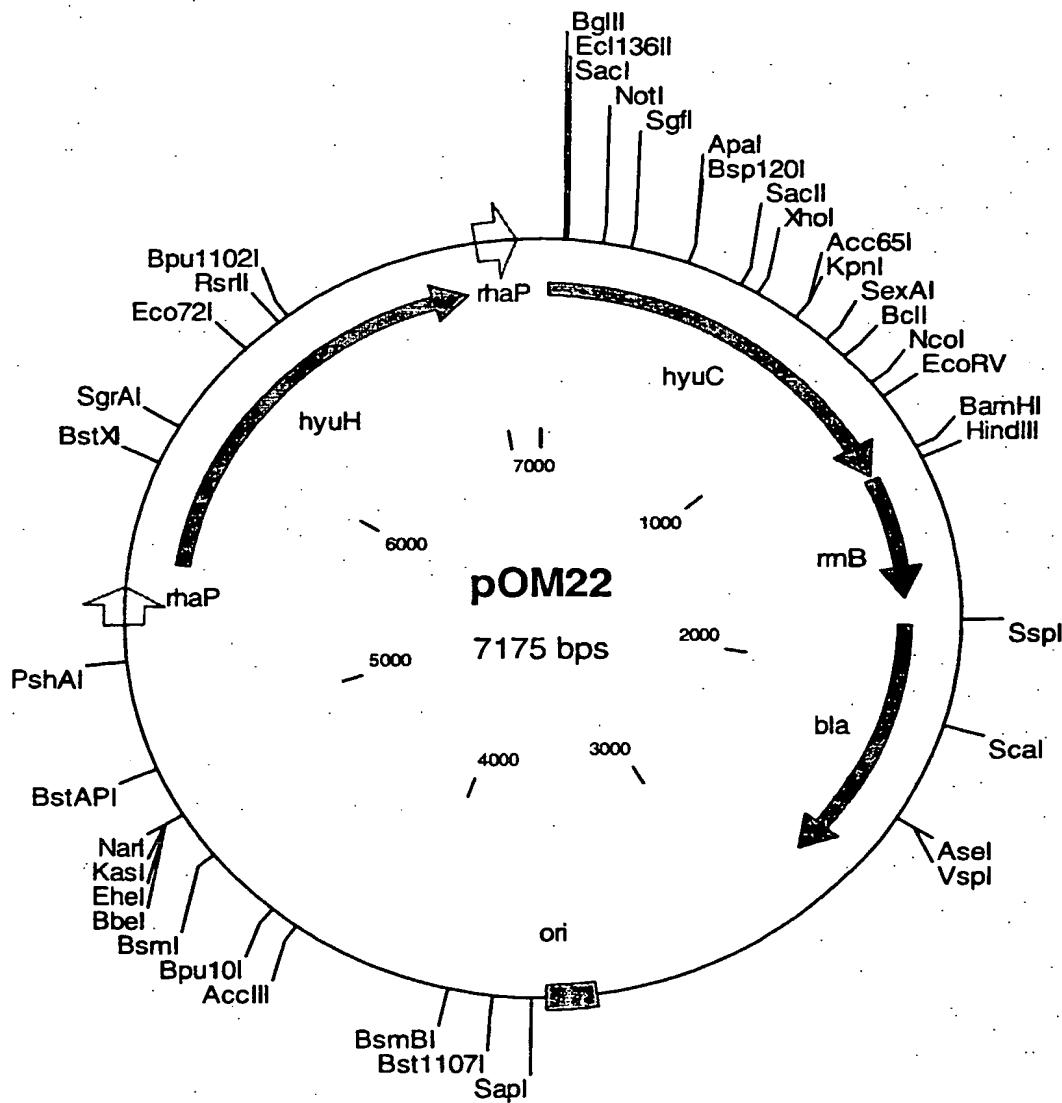
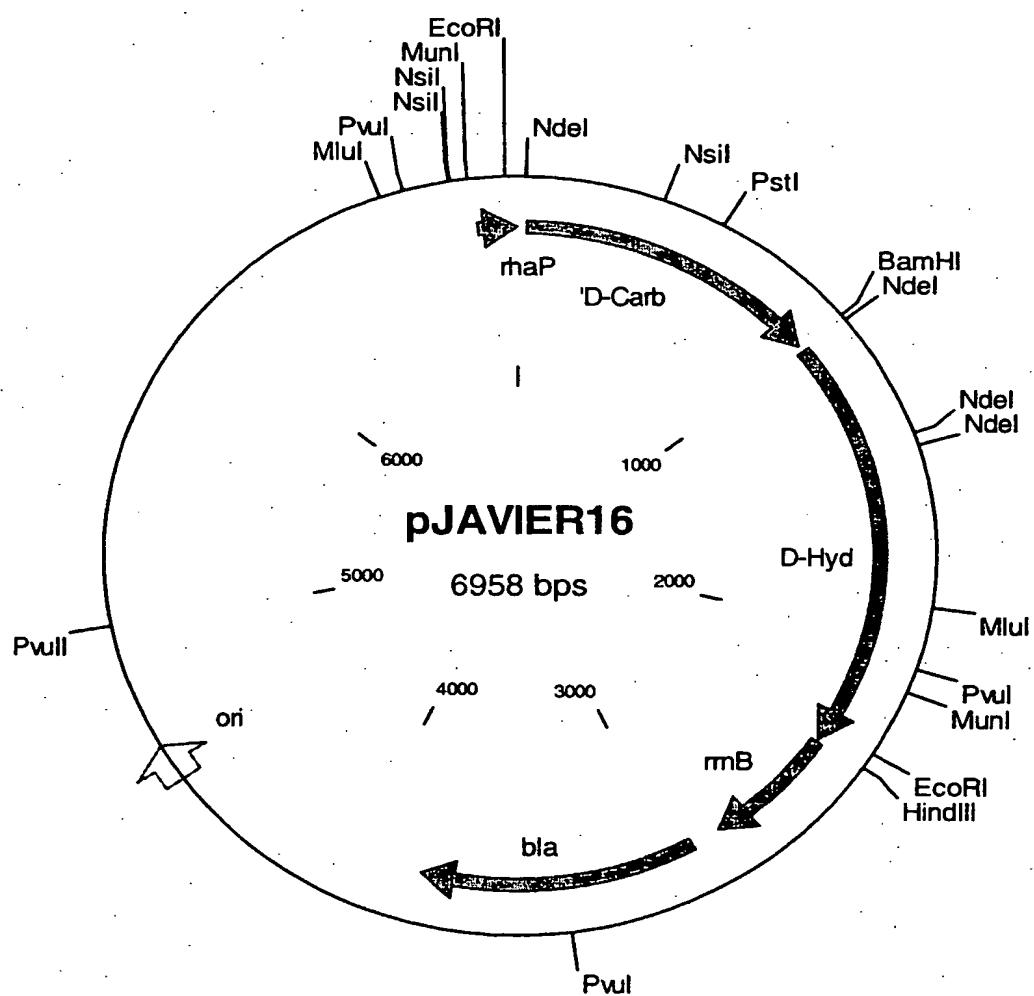


Fig. 4



1 IAP12 Rec'd PCT/PTO 05 DEC 2005

SEQUENCE LISTING

<110> Degussa AG

5 <120> Screening process for hydantoin racemases

<130> 030115 AM

<140>

10 <141>

<160> 16

15 <170> PatentIn Ver. 2.1

<210> 1

<211> 6

<212> PRT

20 <213> Artificial sequence

<220>

<223> Description of the artificial sequence

Consensus sequence

25 <400> 1

Phe Xaa Asp Xaa Gly Leu

1 5

30 <210> 2

<211> 236

<212> PRT

<213> Arthrobacter crystallopoietes

35 <400> 2

Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu
1 5 10 15

40 Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile

20 25 30

Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
35 40 4545 Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
50 55 60Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Gly Asp
65 70 75 8050 Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
85 90 9555 Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
100 105 110Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
115 120 125

Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 130 135 140
 Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 5 145 150 155 160
 Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175
 10 Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 180 185 190
 Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205
 15 Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 210 215 220
 Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 20 225 230 235

<210> 3
 <211> 711
 25 <212> DNA
 <213> Artificial sequence

<220>
 30 <223> Description of the artificial sequence: 1BG7
 <220>
 <221> CDS
 <222> (1)..(711)

35 <400> 3
 atg aga atc ctc gtg atc aac ccc aac agt tcc agc gcc ctt act gaa 48
 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ala Leu Thr Glu
 1 5 10 15

40 tcg gtt gcg gac gca gca caa caa gtt gtc gcg acc ggc acc ata att 96
 Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
 20 25 30

45 tct gcc atc aac ccc tcc aga gga ccc gcc gtc att gaa ggc agc ttt 144
 Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
 35 40 45

50 gac gaa gca ctg gcc acg ttc cat ctc att gaa gag gtg gag cgc gct 192
 Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
 50 55 60

55 gag cgg gaa aac ccg ccc gac gcc tac gtc atc gca tgt ttc agg gat 240
 Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Arg Asp
 65 70 75 80

55 ccg gga ctt gac gcg gtc aag gag ctg act gac agg cca gtg gta gga 288
 Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
 85 90 95

gtt gcc gaa gct gca atc cac atg tct tca ttc gtc gcg gcc acc ttc 336
 Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
 100 105 110

5 tcc att gtc agc atc ctc ccg agg gtc agg aaa cat ctg cac gaa ctg 384
 Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
 115 120 125

10 gta cgg caa gcg ggg gcg acg aat cgc ctc gcc tcc atc aag ctc cca 432
 Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 130 135 140

15 aat ctg ggg gtg atg gcc ttc cat gag gac gaa cat gcc gca ctg gag 480
 Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 145 150 155 160

20 acg ctc aaa caa gcc gcc aag gag gcg gtc cag gag gac ggc gcc gag 528
 Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175

25 tcg ata gtg ctc gga tgc gcc ggc atg gtg ggg ttt gcg cgt caa ctg 576
 Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 180 185 190

30 agc gac gaa ctc ggcc gtc cct gtc atc gac ccc gtc gag gca gct tgc 624
 Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205

35 cgc gtg gcc gag agt ttg gtc gct ctg ggc tac cag acc agc aaa gcg 672
 Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 210 215 220

40 aac tcg tat caa aaa ccg aca gag aag cag tac ctc tag 711
 Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235

<210> 4
 <211> 237
 <212> PRT
 <213> Artificial sequence
 <223> Description of the artificial sequence:1BG7

45 <400> 4
 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu
 1 5 10 15

Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
 20 25 30

50 Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
 35 40 45

Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
 55 55 60

Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Arg Asp
 65 70 75 80

Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
 85 90 95

Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
 5 100 105 110

Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
 115 120 125

10 Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 130 135 140

Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 145 150 155 160

15 Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175

Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 20 180 185 190

Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205

25 Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 210 215 220

Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235

30 <210> 5
 <211> 711
 <212> DNA
 35 <213> Artificial sequence

<220>
 <223> Description of the artificial sequence: 3CH11

40 <220>
 <221> CDS
 <222> (1)...(711)

<400> 5
 45 atg aga atc ctc gtg atc aac ccc aac agt tcc agc gcc ctt act gaa 48
 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ala Leu Thr Glu
 1 5 10 15

50 tcg gtt gcg gac gca gca caa caa gtt gtc gcg acc ggc acc ata att 96
 Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
 20 25 30

tct gcc atc aac ccc tcc aga gga ccc gcc gtc att gaa ggc agc ttt 144
 Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
 55 35 40 45

gac gaa gca ctg gcc acg ttc cat ctc att gaa gag gtg gag cgc gct 192
 Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
 50 55 60

gag cgg gaa aac ccg ccc gac gcc tac gtc atc gca tgt ttc gag gat 240
 Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Glu Asp
 65 70 75 80
 5
 ccg gga ctt gac gcg gtc aag gag ctg act gac agg cca gtc gta gga 288
 Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
 85 90 95
 10 gtt gcc gaa gct gca atc cac atg tct tca ttc gtc gcg gcc acc ttc 336
 Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
 100 105 110
 15 tcc att gtc agc atc ctc ccg agg gtc agg aaa cat ctg cac gaa ctg 384
 Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
 115 120 125
 20 gta cgg caa gcg ggg gcg acg aat cgc ctc gcc tcc atc aag ctc cca 432
 Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 130 135 140
 25 aat ctg ggg gtg atg gcc ttc cat gag gac gaa cat gcc gca ctg gag 480
 Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 145 150 155 160
 30 acg ctc aaa caa gcc gcc aag gag gcg gtc cag gag gac ggc gcc gag 528
 Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175
 35 tcg ata gtg ctc gga tgc gcc ggc atg gtg ggg ttt gcg cgt caa ctg 576
 Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 180 185 190
 40 agc gac gaa ctc ggc gtc cct gtc atc gac ccc gtc gag gca gct tgc 624
 Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205
 45 cgc gtg gcc gag agt ttg gtc gct ctg ggc tac cag acc agc aaa gcg 672
 Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 210 215 220
 aac tcg tat caa aaa ccg aca gag aag cag tac ctc tag 711
 Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235
 50 <210> 6
 <211> 237
 <212> PRT
 <213> Artificial sequence
 <223> Description of the artificial sequence:3CH11
 <400> 6
 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu
 55 1 5 10 15
 Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
 20 25 30

Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
 35 40 45

Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
 5 50 55 60

Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Glu Asp
 65 70 75 80

10 Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
 85 90 95

Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
 100 105 110

15 Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
 115 120 125

Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 20 130 135 140

Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 145 150 155 160

25 Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175

Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 30 180 185 190

Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205

Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 35 210 215 220

Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235

40 <210> 7
 <211> 711
 <212> DNA
 <213> Artificial sequence

45 <220>
 <223> Description of the artificial sequence:AE3

50 <220>
 <221> CDS
 <222> (1)..(711)

<400> 7
 atg aga atc ctc gtg atc aac ccc aac agt tcc agc gcc ctt act gaa 48
 55 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ala Leu Thr Glu
 1 5 10 15

	tcg gtt gcg gac gca gca caa caa gtt gtc gcg acc ggc acc ata att	96
	Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile	
	20 25 30	
5	tct gcc atc aac ccc tcc aga gga ccc gcc gtc att gaa ggc agc ttt	144
	Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe	
	35 40 45	
10	gac gaa gca ctg gcc acg ttc cat ctc att gaa gag gtg gag cgc gct	192
	Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala	
	50 55 60	
15	gag cgg gaa aac ccg ccc gac gcc tac gtc atc gca tgt ttc cag gat	240
	Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Gln Asp	
	65 70 75 80	
	ccg gga ctt gac gcg gtc aag gag ctg act gac agg cca gtg gta gga	288
	Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly	
	85 90 95	
20	gtt gcc gaa gct gca atc cac atg tct tca ttc gtc gcg gcc acc ttc	336
	Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe	
	100 105 110	
25	tcc att gtc agc atc ctc ccg agg gtc agg aaa cat ctg cac gaa ctg	384
	Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu	
	115 120 125	
30	gta cgg caa gcg ggg gcg acg aat cgc ctc gcc tcc atc aag ctc cca	432
	Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro	
	130 135 140	
35	aat ctg ggg gtg atg gcc ttc cat gag gac gaa cat gcc gca ctg gag	480
	Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu	
	145 150 155 160	
	acg ctc aaa caa gcc gcc aag gag gcg gtc cag gag gac ggc gcc gag	528
	Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu	
	165 170 175	
40	tcg ata gtg ctc gga tgc gcc ggc atg gtg ggg ttt gcg cgt caa ctg	576
	Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu	
	180 185 190	
45	agc gac gaa ctc ggc gtc cct gtc atc gac ccc gtc gag gca gct tgc	624
	Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys	
	195 200 205	
50	cgc gtg gcc gag agt ttg gtc gct ctg ggc tac cag acc agc aaa gcg	672
	Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala	
	210 215 220	
	aac tcg tat caa aaa ccg aca gag aag cag tac ctc tag	711
	Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu	
55	225 230 235	

<212> PRT
<213> Artificial sequence
<223> Description of the artificial sequence:AE3

5 <400> 8
 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu
 1 5 10 15
 10 Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
 20 25 30
 Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
 35 40 45
 15 Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
 50 55 60
 Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Gln Asp
 65 70 75 80
 20 Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
 85 90 95
 25 Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
 100 105 110
 Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
 115 120 125
 30 Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 130 135 140
 Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 145 150 155 160
 35 Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175
 40 Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 180 185 190
 Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205
 45 Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 210 215 220
 Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235

```
55 <210> 9  
<211> 711  
<212> DNA  
<213> Artificial sequence
```

<220>
<223> Description of the artificial sequence:BB5

<220>
 <221> CDS
 <222> (1)..(711)

5 <400> 9
 atg aga atc ctc gtg atc aac ccc aac agt tcc agc gcc ctt act gaa 48
 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ala Leu Thr Glu
 1 5 10 15

10 tcg gtt gcg gac gca gca caa caa gtt gtc gcg acc ggc acc ata att 96
 Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
 20 25 30

15 tct gcc atc aac ccc tcc aga gga ccc gcc gtc att gaa ggc agc ttt 144
 Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
 35 40 45

20 gac gaa gca ctg gcc acg ttc cat ctc att gaa gag gtg gag cgc gct 192
 Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
 50 55 60

25 gag cgg gaa aac ccg ccc gac gcc tac gtc atc gca tgt ttc ttg gat 240
 Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Leu Asp
 65 70 75 80

30 ccg gga ctt gac gcg gtc aag gag ctg act gac agg cca gtg gta gga 288
 Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
 85 90 95

35 tcc att gtc agc atc ctc ccg agg gtc agg aaa cat ctg cac gaa ctg 384
 Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
 115 120 125

40 gta cgg caa gcg ggg gcg acg aat cgc ctc gcc tcc atc aag ctc cca 432
 Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 130 135 140

45 aat ctg ggg gtg atg gcc ttc cat gag gac gaa cat gcc gca ctg gag 480
 Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 145 150 155 160

50 acg ctc aaa caa gcc gcc aag gag gcg gtc cag gag gac ggc gcc gag 528
 Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175

55 tcg ata gtg ctc gga tgc gcc ggc atg gtg ggg ttt gcg cgt caa ctg 576
 Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 180 185 190

55 agc gac gaa ctc ggc gtc cct gtc atc gac ccc gtc gag gca gct tgc 624
 Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205

cgc gtg gcc gag agt ttg gtc gct ctg ggc tac cag acc agc aaa gcg 672
 Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 210 215 220
 aac tcg tat caa aaa ccg aca gag aag cag tac ctc tag 711
 Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235
 10 <210> 10
 <211> 237
 <212> PRT
 <213> Artificial sequence
 <223> Description of the artificial sequence:BB5
 15 <400> 10
 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu
 1 5 10 15
 20 Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
 20 25 30
 Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
 35 40 45
 25 Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
 50 55 60
 30 Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Leu Asp
 65 70 75 80
 Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
 85 90 95
 35 Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
 100 105 110
 Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
 115 120 125
 40 Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 130 135 140
 45 Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 145 150 155 160
 Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175
 50 Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 180 185 190
 Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205
 55 Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 210 215 220

Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
225 230 235

5 <210> 11
<211> 25
<212> DNA
<213> Artificial sequence

10 <220>
<223> Description of the artifical sequence: Primer5
<400> 11
gccgcaagga atgggtgcatg catcg 25
15
<210> 12
<211> 30
<212> DNA
20 <213> Artificial sequence
<220>
<223> Description of the artificial sequence: Primer6
25 <400> 12
ggtcagggtgg gtccaccgcg ctactgccgc 30

30 <210> 13
<211> 5777
<212> DNA
<213> Artificial sequence

35 <220>
<223> Description of the artificial sequence: Plasmid pOM21
<400> 13
aattcttaag aaggagatat acatatgaga atcctcgtga tcaaccccaa cagttccagc 60
40 gcccttactg aatcggttgc ggacgcagca caacaagttg tcgcgaccgg caccataatt 120
tctgccatca acccctccag aggacccgccc gtcattgaag gcagcttga cgaagcactg 180
45 gcccacgttcc atctcattga agaggtggag cgcgctgagc gggaaaaccc gcccacgc 240
tacgtcatcg catgtttcgg ggatccggga cttgacgcgg tcaaggagct gactgacagg 300
ccagtggttag gagttgccga agctgcaatc cacatgtctt cattcgtcgc ggccaccc 360
50 tccattgtca gcatcctccc gagggtcagg aaacatctgc acgaactggt acggcaagcg 420
ggggcgacga atcgccctcgc ctccatcaag ctcccaaatac tgggggtgat ggcccttccat 480
55 gaggacgaac atgcccgcact ggagacgctc aaacaagccg ccaaggaggc ggtccaggag 540
gacggcgccg agtcgatagt gctcgatgc gccggcatgg tggggtttgc gctcaactg 600
agcgacgaac tcggcgcccc tgtcatcgac cccgtcgagg cagcttgcgg cgtggccgag 660

agtttggtcg ctctggcta ccagaccagc aaagcgaact cgtatcaaaa accgacagag 720
aagcagtacc tctagctgca gccaagcttc tgtttggcg gatgagagaa gatttcagc 780
5 ctgatacaga ttaaatcaga acgcagaagc ggtctgataa aacagaattt gcctggcgcc 840
agtagcgcgg tggtcccacc tgacccatg ccgaactcag aagtgaaacg ccgtacgccc 900
gatggtagtg tggggtctcc ccatgcgaga gtaggaaact gccaggcatc aaataaaaacg 960
10 aaaggctcag tcgaaagact gggctttcg ttttatctgt tgtttgcgg tgaacgctct 1020
cctgagtagg acaaatccgc cgggagcggg tttgaacggt gcgaagcaac ggcccgagg 1080
15 gtggcgggca ggacgcccgc cataaaactgc caggcatcaa attaaggcaga agggcatct 1140
gacggatggc ctttttgcgt ttctacaaac tctttgttt attttctaa atacattcaa 1200
20 atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat cgtccatcc 1260
gacagcatcg ccagtcacta tggcgtgctg ctagcgtat atgcgttgat gcaatttcta 1320
tgcgccacccg ttctcgaggc actgtccgac cgcttggcc gccgcccagt cctgctcgct 1380
25 tcgctacttg gagccactat cgactacgcg atcatggcga ccacaccgt cctgtggatc 1440
ctctacgccc gacgcacgt ggcggcatc accggcgccca caggtgcgggt tgctggcgcc 1500
30 tatacgccg acatcaccga tggggaaat cgggctcgcc acttcgggct catgagcgct 1560
tgtttcggcg tgggtatggt ggcaggcccc gtggccgggg gactgttggg cgccatctcc 1620
ttgcatgcac cattccttgc ggcggcggtg ctcaacggcc tcaacctact actgggctgc 1680
35 ttcctaattgc aggagtgcga taagggagag cgtcgaccga tgcccttgag agcattcaac 1740
ccagtcagct cttccgggtg ggcgggggc atgactatcg tcgcccact tatgactgtc 1800
40 ttctttatca tgcaactcggtt aggcacagggtt cggcagcgc tctgggtcat tttcgccgag 1860
gaccgcttgc gctggagcgc gacgatgatc ggcctgtcgc ttgcggattt cggaatcttgc 1920
cacggccctcg ctcaaggcctt cgtcactggt cccgccccca aacgtttcg cgagaaggcag 1980
45 gcccattatcg cccgcattggc ggccgacgcg ctgggtacg tcttgcgtggc gttcgacg 2040
cgaggctggc tggccttccc cattatgatt cttctcgctt cccggggcat cggatgccc 2100
50 gcggtgcagg ccatgctgtc caggcaggta gatgacgacc atcaggacca gcttcaagga 2160
tcgctcgccg ctcttaccag cctaacttcg atcactggac cgctgatcgatcacggcatt 2220
tatgcccgcct cggcgagcac atggaacggg ttggcatgga ttgtaggcgc cgcctatac 2280
55 cttgtctgcc tccccgcgtt gcgtcgccgt gcatggagcc gggccaccc gacctgaatg 2340
gaagccggcg gcacctcgct aacggattca ccactccaag aattggagcc aatcaattct 2400
tgcggagaac tgtgaatgcg caaaccaacc cttggcagaa catatccatc gcgtccgcca 2460

tctccagcag ccgcacgcgg cgcatctcg ggccatgggg gtcctggcca cgggtgcgca 2520
5 tgatcgtgct cctgtcggtt aggacccggc taggctggcg gggttgcctt actggtagc 2580
agaatgaatc accgatacgc gagcgaacgt gaagcgactg ctgctgcaaa acgtctgcga 2640
cctgagcaac aacatgaatg gtcttcggtt tccgtgttc gtaaaagtctg gaaacgcgga 2700
10 agtcccctac gtgctgctga agtgccgc aacagagagt ggaaccaacc ggtgatacca 2760
cgatactatg actgagagtc aacgcctatga gcggccctcat ttcttattct gagttacaac 2820
15 agtccgcacc gctgtccggt agtccttcc ggtggcgcg gggcatgact atcgtcgccg 2880
cacttatgac tgtcttcctt atcatgcaac tcgttaggaca ggtgccggca gcgcacaaca 2940
gtccccccgc cacggggcct gccaccatac ccacgcccga acaagcgccc tgcaccatta 3000
20 tgttccggat ctgcatcgca ggatgctgct ggctaccctg tggaaacacct acatctgtat 3060
taacgaagcg ctaaccgttt ttatcaggct ctgggaggca gaataaaatga tcataatgctc 3120
25 aattattacc tccacgggga gagcctgagc aaactggcct caggcatttga agaagcacac 3180
ggtcacactg cttccggtag tcaataaaacc ggtaaaccag caatagacat aagcggctat 3240
ttaacgaccc tgccctgaac cgacgaccgg gtcgaatttgc cttcgaatt tctgccattc 3300
30 atccgcttat tatcaatttgc tcaggcgtt caccaggcgt ttaagggcac caataactgc 3360
cttaaaaaaaaaa ttacgccccg ccctgccact catgcagta ctgtttaat tcattaagca 3420
35 ttctgccac atggaagcca tcacagacgg catgatgaac ctgaatcgcc agcggcatca 3480
gcacccctgtc gccttgcgttaaatatttgc ccatggtgaa aacggggcg aagaagttgt 3540
ccatattggc cacgtttaaa tcaaaaactgg tgaaactcac ccagggatttgc gctgagacga 3600
40 aaaacatatt ctcaataaaac ctttaggaa aataggccag gtttcacccg taacacgcca 3660
catcttgcga atatatgtgt agaaaactgcc ggaaatcgctc gtggatttca ctccagagcg 3720
45 atgaaaacgt ttcaatggaa cgggtgttgc tcatggaaaa cgggtgttgc agggtaaca ctatcccata 3780
tcaccagctc accgtcttcc attgccatac gaattccggta tgagcattca tcaggcgcc 3840
aagaatgtga ataaaggccg gataaaaactt gtgcttattt ttctttacgg tctttaaaaa 3900
50 ggccgtataa tccagctgaa cgggtgttgc ataggtacat tgagcaactg actgaaatgc 3960
ctcaaaaatgt tctttacgt gccattggaa tataatcaacg gtggatatac cagtgatttt 4020
55 tttctccatt ttagcttcct tagctcctga aaatctcgat aactcaaaaa atacgcccgg 4080
tagtgcattt attcatttgc ggtgaaagtt ggaaccttgc acgtgcccgtt caacgtctca 4140
tttccggccaa aagttggccc agggcttccc ggtatcaaca gggacaccag gatttattta 4200

ttctgcgaag tgatcttccg tcacaggtat ttattcggcg caaagtgcgt cgggtgatgc 4260
 tgccaaactta ctgatTTAGT gtatgtatggt gttttgagg tgctccagtg gcttctgttt 4320
 5 ctatcagctg tccctcctgt tcagctactg acgggggtggt gcgtAACGGC aaaAGCACCg 4380
 ccggacatca gcgctAGCGG agtgtatact ggcttactat gttggcactg atgagggtgt 4440
 10 cagtgaaAGTg cttcatgtgg caggagaaaa aaggctgcac cggtgcgta gcagaatatg 4500
 tgatacagga tatattccgc ttccTcgctc actgactcgc tacgctcggt cgttcgactg 4560
 cggcgagcgg aaatggctta cgaacggggc ggagatttcc tggaaAGATgc caggaagata 4620
 15 cttaacaggg aagttagagg gcccggcaa agccgtttt ccataggctc cgccccctg 4680
 acaagcatca cgaatctga cgctcaaATC agtgggtggcg aaACCCGACA ggactataaa 4740
 20 gataccaggc gtttcccctg ggggtccct cgtgcgtct cctgttccctg ccttcggtt 4800
 taccgggtgc attccgtgt tatggcccg tttgtctcat tccacgcctg acactcagtt 4860
 ccggtaggc agtgcgtcc aagctggact gtatgcacga accccccgtt cagtcggacc 4920
 25 gctgcgcctt atccggtaac tatacgcttg agtccaaACCC gaaaAGACat gcaaaAGCAC 4980
 cactggcagc agccactggt aattgattt gaggagttt tagtgcgtc atgcggcgg 5040
 30 taaggctaaa ctgaaaaggac aagtTTTGGT gactgcgtc ctccaaGGCA gttacctcg 5100
 ttcaaaAGAT tggtagctca gagaaccttc gaaaaACCGC cctgcaaggc ggtttttcg 5160
 tttcagagc aagagattac ggcgcagacca aaacgatctc aagaAGATCA tcttattaa 5220
 35 cagataaaat atttcaagat ttcaGTGCAA tttatctctt caaatgttagc acctgaAGTC 5280
 agccccatac gatataAGTT gtaattctca tggtagacag cttatcatcg ataAGCTTA 5340
 40 atgcggtagt ttatcacagt taaattgcta acgcAGTCAG gcaccgtgt a taaaatctaa 5400
 caatgcgtc atcgtcatcc tcggcaccgt caccctggat gctgttaggca taggcttggt 5460
 tatGCCGGA ctggggcc tcttgcggga tttagtcatgc cccgcGCCca ccggaaaggag 5520
 45 ctgactgggt tgaaggctct caaggGCATC ggtcgacgct ctcccttatg cgactcctgc 5580
 attaggaAGC agcccAGTAG taggttgagg cggtagcAGC cggccGCCGc aaggaaATGGT 5640
 50 gcatgcATCG atcaccacaa ttcaGCAAAT tgtgaacatc atcACGTTCA tcttccctg 5700
 gttgccaatg gcccatttc ctgtcagtaa cgagaAGGTC gcaatttcAG gcgctttta 5760
 gactggTCGT aatgaac 5777

55

<210> 14
 <211> 7175
 <212> DNA
 <213> Artificial sequence

<220>

<223> Description of the artificial sequence: Plasmid pOM22

5 <400> 14

aattcttaag aaggagatat acatatgacc ctgcagaaaag cgcaagcgna ggcatttag 60
aaagagatct gggagctctc ccggttctcg gcggaaaggcc ccgggtttac ccggctgacc 120
10 tacactccag agcatgccgc cgccggaa acgctcattt cggctatgga agcggccgct 180
ttgagcgttc gtgaagacgc tctcggaaac atcatcgcc gacgtgaagg cactgatccg 240
15 cagctccctg cgatcgcggc cggttcacac ttcgattctg tccgaaacgg cggatgttc 300
gatggcactg caggcgtggt gtgcgcctt gaggctgccc gggatgtct ggagagcggc 360
tacgtgaatc ggcattccatt tgagttcattc gcgatcgtgg aggaggaagg ggccgcattc 420
20 agcagtggca tttttggcgg ccggccatt gcagggttgg tcggcgcacag ggaactggac 480
tctttggttt atgaggatgg agtgcgttt aggcaggcgg ctactgcctt cggcttgaag 540
25 cccggcgaac tgcaggctgc agccgcctt cggcggacc tgcgtgttt tatcgaacta 600
cacattgaac aaggaccat cctcgagcag gagcaaata agatcggagt tgtgacctcc 660
atcggtggcg ttgcgcatt gcgggttgct gtcaaaggca gaagcgcaca cggccgcaca 720
30 accccatgc acctgcgcgc ggtgcgcctt gtacccgcgtt ctctcatggt gcgggaggc 780
aaccggttcg tcaacgagat cggcgttgc acagtggcta ccgttggcca cctcacagt 840
35 gccccggcg gcccggatcc ggtccgggg gaggtggagt tcacactgga cctgcgttct 900
ccgcatgagg agtcgcctcg ggtgttgc aaccgcattt cggcatggt cggcggaggc 960
gcctcgagg cgggtgtggc tgccgtgtt gatgaatttt tcaatctcag cccgggtgcag 1020
40 ctggcttcata ccatggtgga cggcgttgcg gaagcggcct cggccctgca gttcacgcac 1080
cgggatataca gcagtggggc gggccacgac tcgatgttca tcgcccaggat cacggacgac 1140
45 ggaatggttt tcgttccaag ccgtgctggc cggagccacg ttcccgaaga atggaccat 1200
ttcgatgacc ttgcgaagg aactgaggat gtcctccggg taatgaaggc acttgaccgg 1260
ggatcccatc atcatcatca tcattgactg cagccaaatct tctgtttgg cggatgagag 1320
50 aagattttca gcctgataca gattaaatca gaacgcagaa gccgtctgat aaaacagaat 1380
ttgcctggcg gcagtagcgc ggtggtccca cctgacccca tgccgaactc agaagtggaaa 1440
55 cggcgttagcg ccgatggtag tgtgggtct cccatgcga gagtagggaa ctgcccaggca 1500
tcaaataaaa cgaaaggcgc agtcgaaaga ctggccctt cgtttatct gttgtttgtc 1560
ggtaacgct ctcctgagta ggacaaatcc gcccggagcg gatttgaacg ttgcgaagca 1620

acggccccgga gggtggcggg caggacgccc gccataaaact gccaggcatc aaattaagca 1680
gaaggccatc ctgacggatg gccttttgc gtttctacaa actctttgt ttattttct 1740
5 aaatacattc aaatatgtat ccgctcatga gacaataacc ctgataaaatg cttcaataat 1800
attgaaaaag gaagagtatg agtattcaac atttccgtgt cgcccttatt cccttttg 1860
10 cggcattttg ctttcctgtt tttgctcacc cagaaacgct ggtgaaagta aaagatgctg 1920
aagatcagtt gggtgcacga gtgggttaca tcgaactgga tctcaacagc ggtaagatcc 1980
ttgagagttt tcgccccgaa gaacgtttc caatgatgag cactttaaa gttctgctat 2040
15 gtggcgcggt attatccgt gttgacgccc ggcaagagca actcggtcgc cgcatacact 2100
attctcagaa tgacttgggt gagtactcac cagtcacaga aaagcatctt acggatggca 2160
tgacagtaag agaattatgc agtgctgcca taaccatgag tgataacact gcggccaact 2220
20 tacttctgac aacgatcgga ggaccgaagg agctaaccgc tttttgcac aacatgggg 2280
atcatgtAAC tcgccttgat cgttgggaac cggagctgaa tgaagccata ccaaacgacg 2340
25 agcgtgacac cacgatgcct gtagcaatgg caacaacggt gcgcaaacta ttaactggcg 2400
aactacttac tctagcttcc cggcaacaat taatagactg gatggaggcg gataaagttg 2460
30 caggaccact tctgcgctcg gcccttccgg ctggctgggt tattgctgat aaatctggag 2520
ccggtgagcg tgggtctcgc ggtatcattt cagcactggg gccagatggt aagccctccc 2580
gtatcgtagt tatctacacg acggggagtc aggcaactat ggtgaacga aatagacaga 2640
35 tcgctgagat aggtgcctca ctgattaagg attggtaact gtcagaccaa gtttactcat 2700
atataacttta gattgattta aaacttcatt tttatTTaa aaggatctag gtgaagatcc 2760
40 ttttgataa tctcatgacc aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag 2820
acccctaga aaagatcaaa ggatcttctt gagatccttt tttctgcgc gtaatctgct 2880
gcttgcAAAC aaaaaaACCA ccgcttACCA cggtggtttg tttGCCGgat caagagctac 2940
45 caactcttt tccgaaggta actggcttca gcagagcgca gataccaaat actgtccttc 3000
tagttagcc gtagttagcc caccacttca agaactctgt agcaccgcct acatacctcg 3060
50 ctctgctaat cctgttacca gtggctgctg ccagtggcga taagtcgtgt cttaccgggt 3120
tggactcaag acgatagtta ccggataagg cgcagcggc gggctgaacg gggggttcgt 3180
gcacacagcc cagcttggag cgaacgacact acaccgaact gagataccctt cagcgtgagc 3240
55 tatgagaaag cggccacgctt cccgaaggaa gaaaggcggaa caggtatccg gtaagcggca 3300
gggtcggaac aggagagcgc acgagggagc ttccaggggg aaacgcctgg tatctttata 3360
gtcctgtcgg gtttcgcac ctctgacttg agcgtcgatt tttgtatgc tcgtcagggg 3420

ggcggagcct atggaaaaac gccagcaacg cggcctttt acggttcctg gcctttgct 3480
5 ggcctttgc tcacatgttc tttcctgcgt tatcccctga ttctgtggat aaccgtatta 3540
ccgccttga gtgagctgat accgctcgcc gcagccgaac gaccgagcgc agcgagtcag 3600
tgagcgagga agcggaaagag cgcctgatgc ggtattttct ctttacgcat ctgtgcggta 3660
10 tttcacaccg catatatggt gcactctcag tacaatctgc tctgatgccg catagttaag 3720
ccagtataca ctccgctatc gctacgtgac tgggtcatgg ctgcgcggc acacccgcca 3780
15 acacccgctg acgcgcctg acgggcttgt ctgctccgg catccgctta cagacaagct 3840
gtgaccgtct ccgggagctg catgtgtcag aggtttcac cgtcatcacc gaaacgcgcg 3900
aggcagctgc ggtaaagctc atcagcgtgg tcgtgaagcg attcacagat gtctgcctgt 3960
20 tcatccgcgt ccagctcggt gagtttctcc agaagcgta atgtctggct tctgataaag 4020
cgggccatgt taagggcggt ttttcctgt ttggtcactt gatgcctccg tgtaaggggg 4080
25 aatttctgtt catggggta atgataccga tgaaacgaga gaggatgctc acgatacggg 4140
ttactgatga tgaacatgcc cggttactgg aacgttgtga gggtaaacaatggc 4200
ggatgcggcg ggaccagaga aaaatcactc agggtaatg ccagcgcttc gttaatacag 4260
30 atgttaggtgt tccacagggt agccagcagc atcctgcgt gcagatccgg aacataatgg 4320
tgcagggcgc tgacttccgc gtttccagac ttacgaaac acggaaaccg aagaccattc 4380
atgttggcgc tcaggtcgca gacgtttgc agcagcagtc gttcacgtt cgctcgct 4440
35 tcggtgattc attctgctaa ccagtaaggc aaccccgcca gcctagccgg gtcctcaacg 4500
acaggagcac gatcatgcgc acccggtggcc aggacccaac gctgcccggag atgcgcgcg 4560
40 tgcggctgct ggagatggcg gacgcgttgg atatgttctg ccaagggttg gttgcgcatt 4620
tcacagttct ccgcaagaat tgattggctc caattcttgg agtggtaat ccgttagcga 4680
45 ggtgcgcgg gttccattc aggtcgaggt ggcccgctc catgcaccgc gacgcaacgc 4740
ggggaggcag acaaggtata gggcgccgc tacaatccat gccaaccctt tccatgtgt 4800
cgccgaggcg gcataaatcg ccgtgacgt cagcggtcca gtgatcgaag ttaggctgg 4860
50 aagagccgcg agcgatcattt gaagctgtcc ctgatggctg tcatctaccc gcctggacag 4920
catggcctgc aacgcggca tcccgatgcc gccgaaagcg agaagaatca taatggggaa 4980
55 ggccatccag cctcgctcg cgaacgccag caagacgtag cccagcgctc cggccgcatt 5040
gccggcgata atggcctgct tctcgccaa acgtttggtg gcgggaccag tgacgaaaggc 5100
ttgagcgagg gcgtgcaaga ttccgaatac cgcaagcgcac aggccgatca tcgtcgct 5160

ccagcgaaag cggtcctcgc cgaaaatgac ccagagcgct gccggcacct gtcctacgag 5220
ttgcatgata aagaagacag tcataagtgc ggcgacgata gtcatgcccc gcgcccaccc 5280
5 gaaggagctg actgggttga aggctctcaa gggcatcggt cgacgctctc ccttatgcga 5340
ctcctgcatt aggaagcagc ccagtagtag gttgaggccg ttgagcaccg cccggcaag 5400
10 gaatggtgca tgcacatcgacc accacaattc agcaaattgt gaacatcatc acgttcatct 5460
ttccctggtt gccaatggcc cattttcctg tcaagtaacga gaaggtcgcg aattcaggcg 5520
cttttagac tggtcgtaat gaacaattct taagaaggag atatacatat gtttgacgta 5580
15 atagttaga actgccgtat ggtgtccagc gacggaatca ccgaggcaga cattctggtg 5640
aaagacggca aagtgcgcgc aatcagctcg gacacaagtg atgttgaggc gagccgaacc 5700
20 attgacgcgg gtggcaagtt cgtgtatgcgg ggcgtggcg atgaacatgt gcatatcatc 5760
gacatggatc tgaagaaccg gtatggccgc ttcaactcg attccgagtc tgccggccgtg 5820
ggaggcatca ccaccatctt ttagatgcgg tttacattcc cggccaccac cactttggac 5880
25 gccttcctcg aaaagaagaa gcaggcgaaaa cagcggttga aagttgactt cgcgctctat 5940
ggcggtggag tgccggaaaa cctgccccag atccgcaaaa tgcacgacgc cggcgcagtg 6000
30 ggcttcaagt caatgtatggc agcctcagtt cgggcatgt tcgacgcccgt cagcgacggc 6060
gaactgttcg aaatcttcca ggagatcgca gcctgtgggt cagtcggcgt ggtccatgcc 6120
gagaatggaa cgatcattca agcgctccag aagcagatca aagccgctgg tcgcaaggac 6180
35 atggccgcct acgaggcattc ccaaccagtt ttccaggaga acgaggccat tcagcgtcg 6240
ttactactgc agaaagaagc cggctgtcga ctgattgtgc ttcacgttag caaccctgac 6300
40 ggggtcgagc tgatacatcg ggcgaatcc gagggccagg acgtccactg cgagtcgggt 6360
ccgcagtatc tgaatatcac cacggacgac ggcgaacgaa tcggaccgtt tatgaaggc 6420
gccccccccg tccgctcagc cgagatgaac gtcagattat gggaaacaact tgagaacggg 6480
45 ctcatcgaca cccttgggtc agaccacggc ggacatcctg tcgaggacaa agaaccggc 6540
tggaaaggacg tggaaaggc cggcaacggt ggcgtggcc ttgagacatc cctgcctatg 6600
50 atgctgacca acggagtgaa taaaggcagg ctatccttgg aacgcctcgt cgaggtgatg 6660
tgcgagaaac ctgcgaagct cttggcattc tatccgcaga agggcacgct acaggttgg 6720
tccgacgccc atctgctcat cctcgatctg gatattgaca ccaaagtggc tgcctcgcag 6780
55 ttcggatccc tgcataagta cagccccgttc gacggatgc cggtcacggg tgcaccgggtt 6840
ctgacgatgg tgcgcggaaac ggtgggtggca gagaaggag aagttctggc cgagcaggga 6900
ttcgccactg tgcgtcaccgg tcacgactac gaggcgtcga agtgaggatc tcgacgctct 6960

cccttatgcg actcctgcat taggaagcag cccagtagta ggttgaggcc gttgagcacc 7020
5 gccgcccggaa ggaatggtgc atgcacatcgat caccacaatt cagcaaattt tgaacatcat 7080.
cacgttcatc tttccctggt tgccaatggc ccattttcct gtcagtaacg agaagggtcgc 7140
gaattcaggc gctttttaga ctggtcgtaa tgaac 7175

10 <210> 15
<211> 5989
<212> DNA
<213> Artificial sequence

15 <220>
<223> Description of the artificial sequence: Plasmid pDHYH

20 aattcttaag aaggagatat acatatggat gcaaagctac tgggtggcgg cactattgtt 60
tcctcgaccg gcaaaatccg agccgacgtg ctgattgaaa acggcaaagt cgccgctgtc 120
25 ggcatgctgg acgcccgcac gcccggacaca gttgagcggg ttgactgcga cggcaaatac 180
gtcatgcccgg cgggtatcga cgttcacacc cacatcgact ccccccctcat ggggaccacc 240
accgcccgtatg attttgcgtatcga cggaaacgatt gcagccgcta ccggcggaaac aacgaccatc 300
30 gtcgatttcg gacagcagct cgccggcaag aacctgctgg aatccgcaga cgcgcaccac 360
aaaaaggcgc aggggaaatc cgtcattgtat tacggcttcc atatgtgcgt gacgaacctc 420
tatgacaatt tcgattccca tatggcagaa ctgacacagg acggaatctc cagttcaag 480
35 gtcttcatgg cctaccgcgg aaggcctgatg atcaacgcacg gcgaaactgtt cgacatcctc 540
aaggaggatcg gctccagcgg tgccaaacta tgcgtccacg cagagaacgg cgacgtcatc 600
40 gacaggatcg ccggccgaccc ctacgccccaa ggaaaaaaccg ggcccccggac ccacgagatc 660
gcacgccccgc cggaaatcgga agtcaagca gtcagccggg ccatcaagat ctcccccgtatg 720
45 gccgaggtgc cgctgtatcc cgtcatctt tccaccagg gggccgtcga ggaagtagct 780
gccgcgcaga tgacaggatg gccaatcagc gcccggaaacgt gcacccacta cctgtcgctg 840
agccgggaca tctacgacca gcccggattc gagccggcca aagctgtcct cacaccacccg 900
50 ctgcgcacac aggaacacca ggacgcgttg tggagaggca ttaacaccgg tgcgtcagc 960
gtcgctcaggat ccgaccactg ccccttctgc tttgaggaaa agcagcggat gggggcagat 1020
55 gacttccggc agatccccaa cggcggggccc ggcgtggagc accgaatgct cgtatgtat 1080
gagaccgggtg tcgcggagg aaaaatgacg atcgagaaat tcgtcgaggt gactgccgag 1140
aaccgggaca agcaattcga tatgtacccg aaaaaggaa caattgcacc gggctccgat 1200

gcagacatca tcgtggtcga ccccaacgga acaaccctca tcagtgccga cacccaaaaa 1260
caaaacatgg actacacgct gttcgaaggc ttcaaaaatcc gttgctccat cgaccaggtg 1320
5 ttctcgctg gcgacctgat cagcgtcaaa ggcgaatatg tcggcacccg cggccgcggc 1380
gaattcatca agcggagcgc ttggagccac ccgcagttcg aaaaataaaaa gcttggctgt 1440
10 tttggcggat gagagaagat tttcagcctg atacagatta aatcagaacg cagaagcgg 1500
ctgataaaac agaatttgcc tggcggcagt agcgcggtgg tcccacctga ccccatgcgg 1560
aactcagaag tgaaacgcgg tagcgccgat ggtagtgtgg ggtctccca tgcgagagta 1620
15 gggactgcc aggcataaa taaaacgaaa ggctcagtcg aaagactggg ccttcgttt 1680
tatctgttgt ttgtcggtga acgctctcct gagtaggaca aatccgcgg gagcggattt 1740
20 gaacgttgcg aagcaacggc ccggagggtg gcggcagga cgcccgccat aaactgccag 1800
gcatcaaatt aagcagaagg ccattctgac ggttgcgttc tacaaaactct 1860
tttggttatt tttctaaata cattcaaata tgtatccgct catgagacaa taaccctgat 1920
25 aaatgcttca ataataattga aaaaggaaga gtatgagtat tcaacatttc cgtgcgccc 1980
ttattccctt ttttgcggca ttttgccttc ctgttttgc tcacccagaa acgctggta 2040
30 aagtaaaaga tgctgaagat cagttgggtg cacgagtggg ttacatcgaa ctggatctca 2100
acagcggtaa gatcctttag agtttcgccc ccgaagaacg ttttccaatg atgagactt 2160
ttaaagttct gctatgtggc gcggtattat cccgtgttga cgccgggcaa gagcaactcg 2220
35 gtcggcgcat acactatattc cagaatgact tggttgagta ctcaccagtc acagaaaagc 2280
atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtgata 2340
40 acactgcggc caacttactt ctgacaacga tggaggacc gaaggagcta accgctttt 2400
tgcacaacat gggggatcat gtaactcgcc ttgatcggtt ggaaccggag ctgaatgaag 2460
ccataccaaa cgacgagcgt gacaccacga tgcctgttagc aatggcaaca acgttgcgca 2520
45 aactattaaac tggcgaacta cttactctag cttccggca acaattaata gactggatgg 2580
aggcggataa agttgcagga ccacttctgc gctcgccct tccggctggc tggtttattt 2640
50 ctgataaaatc tggagccggc gagcgtgggt ctcgcggtat cattgcagca ctggggccag 2700
atggtaagcc ctccccgtatc gtagttatct acacgacggg gagtcaggca actatggatg 2760
aacgaaatag acagatcgct gagataggtg cctcactgat taagcattgg taactgtcag 2820
55 accaagttta ctcataatata ctttagattt attaaaaact tcattttaa tttaaaaagga 2880
tcttaggtgaa gatcctttt gataatctca tgacccaaat cccttaacgt gagtttcgt 2940
tccactgagc gtcagacccc gtagaaaaaga tcaaaggatc ttcttgagat ctttttttc 3000

tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcggtg gtttgttgc 3060
5 cgatcaaga gctaccaact cttttccga aggttaactgg cttcagcaga ggcgagatac 3120
caaatactgt cttcttagtg tagccgtagt taggccacca cttcaagaac tctgtacac 3180
cgccctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt ggcgataagt 3240
10 cgtgtttac cgggttggac tcaagacgt agttaccgga taaggcgcag cggtcgggct 3300
gaacgggggg ttcgtgcaca cagccagct tggagcgaac gacctacacc gaactgagat 3360
15 acctacagcg ttagctatga gaaagcgcca cgcttccga agggagaaag ggcgacaggt 3420
atccggtaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca ggggaaacg 3480
cctggtatct ttatagtcct gtcgggttcc gccacctctg acttgagcgt cgattttgt 3540
20 gatgctcgtc aggggggcgg agcctatgga aaaacgccag caacgcggcc ttttacgg 3600
tcctggcctt ttgctggcct tttgctcaca tggtcttcc tgcttatcc cctgattctg 3660
25 tggataaccg tattaccgcc tttgagttagt ctgataccgc tcgcccgcagc cgaacgaccg 3720
agcgcagcga gtcagtgagc gaggaaagcgg aagagcgcct gatgcggtat tttctcctta 3780
cgcatctgtg cggtatttca caccgcatat atggtgcaact ctcagtacaa tctgctctga 3840
30 tgccgcatacg ttaagccagt atacactccg ctatcgctac gtgactgggt catggctgcg 3900
ccccgacacc cggcaacacc cgctgacgcg ccctgacggg cttgtctgct cccggcatcc 3960
gcttacagac aagctgtgac cgtctccggg agctgcattgt gtcagaggtt ttcaccgtca 4020
35 tcaccgaaac ggcgaggca gctgcggtaa agctcatcag cgtggcgtg aagcgattca 4080
cagatgtctg cctgttcatc cgcgtccagc tcgtttagtt tctccagaag cgtaatgtc 4140
40 tggcttctga taaagcgggc catgttaagg gcggttttt cctgtttggc cacttgatgc 4200
ctccgtgtaa ggggaattt ctgttcatgg gggtaatgtat accgatgaaa cgagagagga 4260
45 tgctcacgat acgggttact gatgtgaac atgcccggtt actggAACgt tggagggtt 4320
aacaactggc ggtatggatg cggcgggacc agagaaaaat cactcagggt caatgccagc 4380
gttcgttaa tacagatgtt ggtgttccac agggtagcca gcagcatcct gcgtgcaga 4440
50 tccggaacat aatggtgcaag ggcgctgact tccgcgttcc cagactttac gaaacacgg 4500
aaccgaagac cattcatgtt gttgctcagg tcgcagacgt tttgcagcag cagtcgttc 4560
55 acgttcgctc gctatcggt gattcattct gctaaccagt aaggcaaccc cgccagccta 4620
gccgggtcct caacgacagg agcacgatca tgccgcacccg tggccaggac ccaacgctgc 4680
ccgagatgcg ccgcgtgcgg ctgctggaga tggcggacgc gatggatatg ttctgccaag 4740

ggttggtttg cgcatcaca gttctccgca agaattgatt ggctccaatt cttggagtgg 4800
tgaatccgtt agcgaggtgc cgccggcttc cattcaggc gagggtggccc ggctccatgc 4860
5 accgcgcacgc aacgcgggga ggcagacaag gtatagggcg gcgcctacaa tccatgccaa 4920
cccgttccat gtgctcgccg aggcggcata aatcgccgtg acgatcagcg gtccagtgtat 4980
10 cgaagtttagg ctggtaagag ccgcgagcga tccttgaagc tgcctgtat ggtcgatc 5040
tacctgcctg gacagcatgg cctgcaacgc gggcatcccg atgcccggg aagcgagaag 5100
aatcataatg gggaaaggcca tccagcctcg cgtcgcgaac gccagcaaga cgtagcccg 5160
15 cgcgtcgccg gccatgccgg cgataatggc ctgtttctcg cggaaacgtt tggggcgccg 5220
accagtgacg aaggctttagg cgagggcgtg caagattccg aataccgcaa gcgacaggcc 5280
20 gatcatcgtc gcgctccacgc gaaagcggtc ctcgcccggaa atgacccaga ggcgtcgccg 5340
cacctgtcct acgagttgca tgataaaagaa gacagtcata agtgcggcga cgatagtcat 5400
gccccgcgcc caccggaaagg agctgactgg gttgaaggct ctcaaggcca tcggtcgacg 5460
25 ctctccctta tgcgactcct gcatttagaa gcagccagt agtaggttga ggccgttgag 5520
caccggccgc gcaaggaatg gtgcattgctc gatggctacg agggcagaca gtaagtggat 5580
30 ttaccataat cccttaattt tacgcaccgc taaaacgcgt tcagcgcgtat cacggcagca 5640
gacaggtaaa aatggcaaca aaccacccta aaaactgcgc gatcgcgcct gataaaatttt 5700
aaccgtatga atacctatgc aaccagaggg tacaggccac attacccca cttaatccac 5760
35 tgaagctgcc attttcatg gtttcacccat cccagcgaag ggccatgcgt gcatcgaaat 5820
taatacgacg aaattaatac gactcaactat agggcaattt cgatcaccac aattcagcaa 5880
40 attgtgaaca tcatcacgtt catcttccc tggttgccaa tggcccatatt tcctgtcagt 5940
aacgagaagg tcgcgaattt aggcgtttt tagactggtc gtaatgaac 5989

45 <210> 16
<211> 6958
<212> DNA
<213> Artificial sequence

50 <220>
<223> Description of the artificial sequence: Plasmid
pJAVIER16

55 <400> 16
aattcttaag aaggagatat acatatggcg aaaaacttga tgctcgccgt cgctcaagtc 60
ggcggtatcg atagttcgga atcaagaccc gaagtcgtcg cccgcttgcat tgccctgctg 120
gaagaagcag cttcccaggg cgccgaactg gtggctttc ccgaactcac gctgaccacg 180

ttcttccgc gtacctggtt cgaagaaggc gacttcgagg aataacttcga taaatccatg 240
cccaatgacg acgtcgccgc cctttcgaa cgcgccaaag accttggcgt gggcttctac 300
5 ctcggatacg cggaactgac cagtgatgag aagcggtaca acacatcaat tctggtaac 360
aagcacggcg acatcgtcgg caagtaccgc aagatgcac tggcgggcca cgccgataac 420
cgggaaaggac taccgaacca gcaccttcaa aagaaatact tccgcgaagg agatctcga 480
10 ttcgggtgtct tcgacttcca cggcgtgcag gtcggaatgt gtctctgcaa cgaccggcga 540
tggccggagg tctaccgctc tttggccctg cagggagcag agtcgtcgt cctggctac 600
15 aacaccccg atttcgttcc cggctggcag gaagagcctc acgcgaagat gttcacgcac 660
cttcttcac ttcaggcagg ggcataccag aactcggtat ttgtggcggc tgccggcaag 720
20 tcgggcttcg aagacgggca ccacatgatc ggcggatcag cggtcgccgc gcccagcggc 780
gaaatcctgg caaaagcagc cggcgaggc gatgaagtgc tcgttgtcaa agcagacatc 840
gacatggcga agccctataa ggaagcgtc ttgcacttcg ccgcacatcg ggcggccgac 900
25 gcatacggca tcatcgccga aaggaaaggg cggggcgc cactgcccgt cccgttcaac 960
gtgaatgact aaggatccga aggagatata catatggatc caaagctact gttggcggc 1020
actattgttt cctcgaccgg caaaatccga gccgacgtgc tgattgaaaa cggcaaagtc 1080
30 gccgctgtcg gcatgctgga cggcgcgacg cggacacag ttgagcgggt tgactgcac 1140
ggcaaatacg tcatgcccgg cggtatcgcac gttcacaccc acatgcactc cccctcatg 1200
35 gggaccacca cgcgcgatga tttgtcagc ggaacgattt cagccgtac cggcggaaaca 1260
acgaccatcg tcgatttcgg acagcagctc gccggcaaga acctgctgga atccgcagac 1320
40 ggcgcaccaca aaaaggcgc gggaaatcc gtcattgatt acggcttcca tatgtgcgtg 1380
acgaacctct atgacaattt cgatccccat atggcagaac tgacacagga cggaaatctcc 1440
agttcaagg tcttcatggc ctaccgcgaa agcctgatga tcaacgacgg cgaactgttc 1500
45 gacatcctca agggagtcgg ctccagcggt gccaaactat gcgtccacgc agagaacggc 1560
gacgtcattcg acaggatcgc cggcgcaccc tacgccccaa gaaaaaccgg gcccgggacc 1620
50 cacgagatcg cacgcccccc ggaatcgaa gtcgaagcag tcagccggc catcaagatc 1680
tccccggatgg ccgaggtgcc gctgtatttc gtgcatttt ccacccaggg gggcgtcgag 1740
gaagtagctg cgcgcagat gacaggatgg ccaatcagcg cggaaacgtg cacccactac 1800
55 ctgtcgctga gcccggacat ctacgaccag cgggattcg agccggccaa agctgcctc 1860
acaccaccgc tgcgcacaca ggaacaccag gacgcgttgt ggagaggcat taacaccgg 1920
gacgtcagcg tcgtcagttc cgaccactgc cccttctgtc ttgaggaaaa gacgcggatg 1980

ggggcagatg acttccggca gatccccaac ggccggcccg gcgtggagca ccgaatgctc 2040
5 gtatgtatg agaccggtgt cgccgaagga aaaatgacga tcgagaaatt cgtcgaggtg 2100
actgccgaga acccggccaa gcaattcgat atgtacccga aaaagggAAC aattgcacccg 2160
ggctccgatg cagacatcat cgtggtcgac cccaacggaa caaccctcat cagtgccgac 2220
10 accaaaaaac aaaacatgga ctacacgctg ttcgaaggct tcaaaatccg ttgctccatc 2280
gaccagggtgt tctcgcgtgg cgacctgatc agcgtcaaag gccaatatgt cggcacccgc 2340
ggccgcggcg aattcatcaa gcggagcgct tggagccacc cgcagttcga aaaataaaaag 2400
15 cttggctgtt ttggcggatg agagaagatt ttcagcctga tacagattaa atcagaacgc 2460
agaagcggtc tgataaaaaca gaatttgccct ggccgcgatc ggcgcgtgtt cccacctgac 2520
20 cccatgccga actcagaagt gaaacgccgt agcgccgatg gtagtgtggg gtctccccat 2580
gcgagagtag ggaactgccca ggcataatc aaaacgaaag gctcagtgcg aagactgggc 2640
25 ctttcgtttt atctgttgcgaa cgctctccctg agtaggacaa atccgcggg 2700
agcggatttg aacgttgcga agcaacggcc cggagggtgg cggcaggac gcccgcata 2760
aactgccagg catcaaaatta agcagaaggc catcctgacg gatggcctt ttgcgtttct 2820
30 acaaacttt ttgttattt ttctaaatac attcaaataat gtatccgctc atgagacaat 2880
aacctgata aatgcttcaa taatattgaa aaaggaagag tatgagtatt caacatttcc 2940
gtgtccctt tattccctt tttgcggcat tttgccttcc tgttttgtt caccagaaa 3000
35 cgctggtgaa agtaaaagat gctgaagatc agttgggtgc acgagtgggt tacatcgaac 3060
tggatctcaa cagcggtaag atccttgaga gtttcgccc cgaagaacgt tttccaatga 3120
40 tgagcacttt taaagttctg ctatgtggcg cggattttatc ccgtgttgac gcccggcaag 3180
agcaactcgg tcgcccata cactatttctc agaatgactt ggttgagtac tcaccagtca 3240
45 cagaaaagca tcttacggat ggcataacatc taagagaatt atgcagtgtt gccataacca 3300
tgagtgataa cactgcggcc aacttacttc tgacaacatc cggaggaccc aaggagctaa 3360
ccgctttttt gcacaacatc gggatcatc taactcgcct tgatcggtgg gaaccggagc 3420
50 tgaatgaagc cataccaaac gacgagcgtg acaccacatc gcctgtatc atggcaacaa 3480
cgttgcgcaa actattaact ggcgaactac ttactctac ttcccgccaa caattaatag 3540
actggatgga ggcggataaa gttgcaggac cacttctgcg ctcggccctt ccggctggct 3600
55 ggtttattgc tgataaatct ggagccgggtg agcgtggtc tcgcggatc attgcagcac 3660
tggggccaga tggtaagccc tcccgatc tagttatcta cacgacgggg agtcaggcaa 3720

ctatggatga acgaaataga cagatcgctg agataggtgc ctcactgatt aagcattgg 3780
aactgtcaga ccaagttac tcataatac tttagattga tttaaaactt cattttaa 3840
5 ttaaaaggat ctaggtgaag atccttttgcataatctcat gacaaaatc ccttaacgtg 3900
agtttcgtt ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc 3960
10 cttttttct gcgcgtatc tgctgcttgc aaacaaaaaa accaccgcta ccagcggtgg 4020
tttggggcc ggtcaagag ctaccaactc ttttccgaa ggtaactggc ttcagcagag 4080
cgccagatacc aaataactgtc cttctagtgt agccgttagtt aggcaccac ttcaagaact 4140
15 ctgttagcacc gcctacatac ctcgctctgc taatccgtt accagtggct gctgccagtg 4200
gcgataagtc gtgtcttacc ggggtggact caagacgata gttaccggat aaggcgcagc 4260
ggtcgggctg aacgggggggt tcgtgcacac agcccgctt ggagcgaacg acctacaccg 4320
20 aactgagata cctacagcgt gagctatgag aaagcgccac gcttccgaa gggagaaagg 4380
cgacagggta tccggtaagc ggcagggtcg gaacaggaga ggcgcacgagg gagcttccag 4440
25 ggggaaacgc ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcg 4500
gatttttgtg atgctcgtaa gggggcgga gcctatggaa aaacgcccagc aacgcggcct 4560
30 ttttacgggtt cctggccctt tgctggcctt ttgctcacat gtttttcct gcgttatccc 4620
ctgattctgt ggataaccgt attaccgcct ttgagtggc tgataaccgt cgcgcagcc 4680
gaacgaccga ggcgcagcggag tcagtggcgg aggaagcgga agagcgccctg atgcggatt 4740
35 ttctccttac gcatctgtgc ggtatttac accgcataata tggtcactc tcagtacaat 4800
ctgctctgat gcccatagt taagccagta tacactccgc tatcgctacg tgactggc 4860
atggctgcgc cccgacaccc gccaacaccc gctgacgcgc cctgacgggc ttgtctgctc 4920
40 ccggcatccg cttacagaca agctgtgacc gtctccggga gctgcatttg tcagaggtt 4980
tcaccgtcat caccgaaacg cgcgaggcag ctgcggtaaa gctcatcagc gtggcgtga 5040
45 agcgattcac agatgtctgc ctgttcatcc gcgtccagct cggttagtt ctccagaagc 5100
gttaatgtct ggcttctgat aaagcgccc atgttaaggg cggtttttc ctgtttggc 5160
acttgatgcc tccgtgtaa gggaaatttc tgttcatggg ggtaatgata ccgatgaaac 5220
50 gagagaggat gctcacgata cgggttactg atgatgaaca tgcccggtta ctggAACgtt 5280
gtgagggtaa acaactggcg gtatggatgc ggcgggacca gagaaaaatc actcagggtc 5340
55 aatgccagcg ctgcgttaat acagatgtag gtgttccaca ggtagccag cagcatcctg 5400
cgatgcagat ccggaaacata atggcgcagg gcgcgtactt ccgcgtttcc agactttacg 5460
aaacacggaa accgaagacc attcatgttg ttgctcaggt cgccagacgtt ttgcagcagc 5520

agtcgcttca cgttcgctcg cgtatcggtg attcattctg ctaaccagta aggcaacccc 5580
5 gccagcctag ccgggtcctc aacgacagga gcacgatcat gcgccacccgt ggcaggacc 5640
caacgctgcc cgagatgcgc cgctgccc tgctggagat ggcggacgcg atggatatgt 5700
tctgccaagg gttggtttgc gcattcacag ttctccgcaa gaattgattt gctccaattc 5760
10 ttggagtggt gaatccgtta gcgaggtgcc gccggcttcc attcaggctcg aggtggcccg 5820
gctccatgca cccgcacgca acgcggggag gcagacaagg tatagggccg cgcctacaat 5880
15 ccatgccaac ccgttccatg tgctgcccga ggcggcataa atcgccgtga cgatcagcgg 5940
tccagtgtatc gaagtttaggc tggtaagagc cgcgagcgat ctttgaagct gtccctgtatg 6000
gtcgtcatct acctgcctgg acagcatggc ctgcaacgcg ggcattccga tgccgcccga 6060
20 agcgagaaga atcataatgg ggaaggccat ccagcctcgc gtcgcgaacg ccagcaagac 6120
gtagcccgacgc gcgtcgcccg ccatgcccgc gataatggcc tgcttctcgc cgaaacgttt 6180
25 ggtggccggga ccagtgacga aggcttgagc gagggcgtgc aagattccga ataccgcaag 6240
cgacaggccg atcatcgctcg cgctccagcg aaagcggtcc tcgcccggaaa tgacccagag 6300
cgctgcccggc acctgtcccta cgagttgcat gataaagaag acagtcataa gtgcggcgac 6360
30 gatagtcatg ccccgcccc accggaagga gctgactggg ttgaaggctc tcaagggcat 6420
cggtcgacgc tctcccttat gcgactcctg cattaggaag cagccagta gtaggttag 6480
35 gccgttgagc accgcccggc caaggaatgg tgcatgctcg atggctacga gggcagacag 6540
taagtggatt taccataatc ccttaattgt acgcaccgct aaaacgcgtt cagcgcgatc 6600
acggcagcag acaggtaaaa atggcaacaa accaccctaa aaactgcgcg atcgccctg 6660
40 ataaatttta accgtatgaa tacctatgca accagagggt acaggccaca ttacccac 6720
ttaatccact gaagctgcca ttttcatgg tttcaccatc ccagcgaagg gcatgcgt 6780
45 catcgaaatt aatacgacga aattaatacg actcactata gggcaattgc gatcaccaca 6840
attcagcaaa ttgtgaacat catcacgttc atctttccct ggttgccaaat ggccatttt 6900
cctgtcagta acgagaaggt cgcaattca ggccgtttt agactggctcg taatgaac 6958
50

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.